

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2002, 09:38:05 ; Search time 31.55 Seconds  
(without alignments)  
1559.610 Million cell updates/sec

Title: US-09-829-936a-22  
Perfect score: 2513  
Sequence: 1 MLPACSLPGLLLWALLLL.....MSYRASSVRLTVFGAYTF 443

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802:\*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2505	99.7	443	AAW32110	Human extracellular
2	2505	99.7	443	AA116587	Extracellular prote
3	2505	99.7	443	AAAB33418	Human PRO226 prote
4	2505	99.7	443	AA184707	A human p53 mutant
5	2505	99.7	443	AA155850	Human S1-5 ECOMP-li
6	2505	99.7	443	AAU12330	Human PRO226 polyp
7	2495	99.3	443	AA192533	Human protein sequ
8	2392	95.2	443	AA184706	Amino acid sequenc
9	2208	87.9	433	AA184706	Lung cancer associ
10	1647	65.5	295	AA184708	C-terminal of p53
11	1619	64.4	295	AA184705	The C-terminal of

12	1518	60.4	274	21	AA176081	Human EGF extracel
13	1518	60.4	274	22	AA156020	Skin cell protein,
14	1379	54.9	493	22	AA172892	Human EFEMP1. Hom
15	1379	54.9	493	22	AA148077	Human extracellular
16	1283	51.1	448	19	AA179739	Human EGF protein
17	1283	51.1	448	20	AA195709	Homo sapiens fetal
18	1283	51.1	448	20	AA194281	Human extracellular
19	1283	51.1	448	21	AA157058	Amino acid sequenc
20	1283	51.1	448	21	AA154989	Full length human
21	1283	51.1	448	22	AA193573	Human polypeptide,
22	1277	50.8	448	20	AA108063	Human EGF-like hom
23	1277	50.8	448	22	AA109227	Human PRO polypept
24	1277	50.8	448	22	AA131183	Amino acid sequenc
25	1269	50.5	448	21	AA156750	Smooth muscle prol
26	1269	50.5	448	21	AA154990	Full length mouse
27	1262	50.2	461	21	AA156752	Smooth muscle prol
28	1262	50.2	461	21	AA154991	Full length mouse
29	1258.5	50.1	423	21	AA156751	Smooth muscle prol
30	1258.5	50.1	423	21	AA156753	Smooth muscle prol
31	1165.5	46.4	392	18	AA131705	Human extracellular
32	1147	45.6	387	22	AA168188	Extracellular prot
33	1143.5	45.5	350	20	AA108066	Human EGF-like pro
34	1074.5	42.8	335	21	AA176008	Rat EGF extracellu
35	1074.5	42.8	335	22	AA155947	Human EGF extracellu
36	1052	41.9	413	22	AA180391	Skin cell protein,
37	1052	41.9	451	22	AA180440	Secreted protein e
38	905	36.0	348	22	AA108490	Gene #21 associate
39	797.5	31.7	576	22	AA180174	Human EGF-like pro
40	797.5	31.7	576	22	AA180175	Human protein SEQ
41	797.5	31.7	576	22	AA141932	Human polypeptide
42	797.5	31.7	603	22	AA179191	Human protein SEQ
43	797.5	31.7	650	22	AA179190	Human protein SEQ
44	791.5	31.5	636	22	AA140146	Human polypeptide
45	711	28.3	683	12	AA111150	Fibulin C. Homo s

#### ALIGNMENTS

RESULT 1  
AAW32110  
ID AAW32110 standard; Protein; 443 AA.  
XX  
AC AAW32110;  
XX  
DT 14-APR-1998 (first entry)  
XX  
DE Human extracellular/epidermal growth factor HCABA58X.  
XX  
KW Extracellular/epidermal growth factor; HCABA58X; human; ss.  
KW vascular smooth muscle proliferation; Marfan syndrome; dementia;  
KW wound healing; alopecia; neurological disorder; ocular disorder;  
KW kidney disorder; liver disorder; embryogenesis; angiogenesis;  
KW antagonist; corneal inflammation; psoriasis; diabetes; therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO9738012-A1.  
XX  
PD 16-OCT-1997.  
XX  
PF 10-APR-1996; 96WO-US05033.  
XX  
PR 10-APR-1996; 96WO-US05033.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Olsen HS, Ruben SM;  
XX  
DR WPI; 1997-512646/47.  
XX  
N-PSDB; AA188974.  
XX  
PT DNA encoding extracellular-epidermal growth factor HCABA58X - useful

PT for treatment and diagnosis of e.g. wounds, neurological disease,  
XX neoplasia, psoriasis etc.  
PS Claim 11; Fig 1; 47pp; English.  
XX This human polypeptide, designated HCABA58X, was identified on the  
CC basis of homology as an extracellular protein-like/epidermal growth  
CC factor-like protein. Its amino acid sequence was deduced from a  
CC cDNA clone (see AAT88974) isolated from an osteoclastoma cDNA library,  
CC and shows 51% identity and 30% similarity to human extracellular  
CC protein. Recombinant HCABA58X polypeptides (the polypeptide  
CC comprising amino acids 1-419 is also claimed) can be expressed in  
CC bacterial, insect, mammalian or plant cells. The polypeptides, and  
CC polynucleotides encoding them, can be used e.g. to induce DNA  
CC synthesis, to regulate vascular smooth muscle proliferation, to  
CC treat Marfan syndrome, to stimulate wound healing, to restore normal  
CC neurological function after trauma or AIDS dementia, to treat ocular  
CC disorders, to treat kidney and liver disorders, to promote hair  
CC follicular development, to stimulate growth and differentiation of  
CC epidermal and epithelial cells in vivo and in vitro, for the  
CC treatment of burns, ulcers and corneal incisions, and to stimulate  
CC embryogenesis and angiogenesis. They can also be used to identify  
CC antagonists (used e.g. to treat corneal inflammation, neoplasia,  
CC tumours, cancers and psoriasis) and agonists, and to raise  
CC diagnostic antibodies.  
XX  
SQ Sequence 443 AA;  
  
Query Match 99.7%; Score 2505; DB 18; Length 443;  
Best Local Similarity 99.8%; Pred. No. 5.6e-135;  
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MLPASCCLPGSLLLWALLLLLLGASPDSEEPDSTECTDGYEWDPSDCHCRDVECLT 60  
Db 1 mlpcascclpgslillwalllllllgaspqdeepsyctcdgyewpdsqhcdrvneclt 60  
  
QY 61 IPEACKGEMKINHGYGCLPRSAAVINDLHGEGPPPPVPPAQHPNCPGPGYEPDDQDS 120  
Db 61 ipeackgemkcinhgygylclprsaavindlhgegppppvppaqhpnpcpgpyepddqds 120  
  
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCQHRVCNL 180  
Db 121 cvdvdecaqalhdcrpsqdcnlnpgsyqctcpdgyrkigpecvdiidcryrycqhrvcnl 180  
  
QY 181 PGSPRCQCEPGFQGLPNRRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHGYYELHRDGF 240  
Db 181 pgspircqcepgfqlgnrrscvdrvnecdmgapceqrcfnsgygtflcrchgyelhrdgs 240  
  
QY 241 CSDIDECSSYLYCQYRCVNEPGRFSCPCPGYQLLATRLCQDIDECESGAHQSEAQTC 300  
Db 241 csdidecsysylcqyrcvnepgrfscpcpgyqllatrlcqdidecesgahqcsaqt 300  
  
QY 301 VNFHGGYRCVDTNRCVPEYIQVSENRLCPASNPCLREOPSSIVHRYMTITTSERSVPADV 360  
Db 301 vnfhgyrcvdtncrvcpeyiqvsenrlcpasnplcreopssivhrymtitstersvpadv 360  
  
QY 361 FQIQATSVYFGAYNAFQIRAGNSQGDYFIRQINNFMFLARVPTGPREYVLDLEMVTM 420  
Db 361 fqiqatsvypgaynafqiragnsggdgyfirqinnvsmflarvptgpreyvidlemvtm 420  
  
QY 421 NSLMSYRASSVLRITVFGAYTF 443  
Db 421 nslmsyrassvrlritvfgaytf 443  
  
RESULT 2  
ID AAY16587  
XX AAY16587 standard; Protein; 443 AA.  
AC AAY16587;  
XX  
DT 23-AUG-1999 (first entry)

XX Extracellular protein-like/Epidermal Growth Factor-like protein.  
DE Human; extracellular protein-like; Epidermal Growth Factor-like protein;  
XX HCABA58X; human osteoclastoma; ligand; EGF receptor; regulation;  
KW vascular smooth muscle cell proliferation; Marfan syndrome;  
KW wound healing; trauma; AIDS dementia; ocular disorder; kidney disorder;  
KW liver disorder; hair follicular development; cell growth; burn; ulcer;  
KW corneal incision.  
OS Homo sapiens.  
XX US5916769-A.  
PN 29-JUN-1999.  
XX 11-APR-1997; 97US-0833963.  
XX 11-APR-1997; 97US-0833963.  
PR 10-APR-1996; 96WO-US05033.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Olsen HS, Ruben SM;  
PI WPI; 1999-394207/33.  
XX DR N-PSDB; AAX60351.  
XX  
PT New extracellular/epidermal growth factor useful for diagnostic and  
PT therapeutic purposes  
XX  
PS Claim 1; Fig 1A-E; 26pp; English.  
XX  
CC The present sequence represents a human extracellular protein-like/  
CC Epidermal Growth Factor-like protein (HCABA58X). The protein is obtained  
CC from a human osteoclastoma cDNA library. Fragments of the full length  
CC HCABA58X gene, at least 15 base long, can be used as hybridization probes  
CC to isolate the full length gene from a cDNA library and to isolate other  
CC genes which have a high sequence similarity or similar biological  
CC activity. The HCABA58X polypeptide can be used as a potential ligand  
CC for an EGF receptor and can be used for the identification,  
CC characterization and cloning of these receptors and to identify new  
CC EGF-type and HCABA58X receptors. Antagonist compounds for the HCABA58X  
CC polypeptides can also be identified using cells that express HCABA58X.  
CC HCABA58X polypeptides can be used to regulate vascular smooth muscle  
CC cell proliferation, treat Marfan syndrome, stimulate wound healing,  
CC restore normal neurological functioning after trauma or AIDS dementia,  
CC to treat ocular disorders, kidney and liver disorders, promote hair  
CC follicular development, stimulate growth and development or epidermal  
CC and epithelial cells in vivo and in vitro and to treat burns, ulcers  
CC and corneal incisions. HCABA58X or its soluble forms can be coupled to  
CC toxic molecules which can then be targeted to specific cells so the  
CC growth factor toxic fusions kill the target cells.  
XX  
SQ Sequence 443 AA;  
  
Query Match 99.7%; Score 2505; DB 20; Length 443;  
Best Local Similarity 99.8%; Pred. No. 5.6e-135;  
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MLPASCCLPGSLLLWALLLLLLGASPDSEEPDSTECTDGYEWDPSDCHCRDVECLT 60  
Db 1 mlpcascclpgslillwalllllllgaspqdeepsyctcdgyewpdsqhcdrvneclt 60  
  
QY 61 IPEACKGEMKINHGYGCLPRSAAVINDLHGEGPPPPVPPAQHPNCPGPGYEPDDQDS 120  
Db 61 ipeackgemkcinhgygylclprsaavindlhgegppppvppaqhpnpcpgpyepddqds 120  
  
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCQHRVCNL 180  
Db 121 cvdvdecaqalhdcrpsqdcnlnpgsyqctcpdgyrkigpecvdiidcryrycqhrvcnl 180

QY 181 PGSFRCCQEPGFGOLGNNSCDVNECDMCAPEQRCFNSYCTFLCRCHQGYELHRDGFs 240  
 Db 181 pgsfrccqepgfgolgnnscdvnecdmgapcegrcfnsygtflcrchqgyelhrdgrfs 240  
 QY 241 CSIDIDCSYSSYLCOYRCVNEPGRFSCHPQGYOLLATRLCQDIDECESGAHOCSEAQTC 300  
 Db 241 csididcsyssylcqvrcvnepgrfscchpqqgyollatrlcqdidcesgahqcseaqt 300  
 QY 301 VNFHGGYRCVDTNRCVPEYIQVSENRCPLCPASNPCLREQPSSIVHRYMTITTSERSYPADV 360  
 Db 301 vnfhggycvdtncrvpeyiqvsenrcplcpasnplcreqpssivhrymtittsersypadv 360  
 QY 361 FOIQATSVYPGANAOIRAGNSOGFYIRQINNVMFLVLRPVTCPREYVLDLEWVTM 420  
 Db 361 fqiqtatvypganaoiragnsogfyirqinnvsmflvrlpvtpreyvldlewmvtm 420  
 QY 421 NSLMSYRASSVLRILTFFVGAYTF 443  
 Db 421 nslmsyrassvrlrtvfgaytf 443

RESULT 3  
 AAB33418  
 ID AAB33418 standard; Protein; 443 AA.  
 AC AAB33418;  
 XX  
 DT 29-JAN-2001 (first entry)  
 XX  
 DE Human PRO226 protein UNQ200 SEQ ID NO:21.  
 KW Human; immune related disease; diagnosis; antiinflammatory; cardiant;  
 KW dermatological; antiarthritic; antirheumatic; immunosuppressive;  
 KW haemostatic; antithyroid; antidiabetic; nootropic; neuroprotective;  
 KW antianaemic; hepatotropic; viricide; antipsoriatic; antiallergic;  
 KW antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;  
 KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;  
 KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;  
 KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;  
 KW autoimmune thrombocytopenia; immune-mediated renal disease;  
 KW demyelinating disease; hepatobiliary disease; Whipple's disease;  
 KW inflammatory bowel disease; gluten-sensitive enteropathy;  
 KW autoimmune disease; immune-mediated skin disease; allergic disease;  
 KW immunological disease; transplantation associated disease;  
 KW graft rejection; graft-versus-host-disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200053758-A2.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF 02-MAR-2000; 2000WO-US05841.  
 XX  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 10-MAR-1999; 99US-0123618.  
 PR 12-MAR-1999; 99US-0123957.  
 PR 23-MAR-1999; 99US-0125775.  
 PR 12-APR-1999; 99US-0128849.  
 PR 20-APR-1999; 99WO-US08615.  
 PR 28-APR-1999; 99US-0131445.  
 PR 04-MAY-1999; 99US-0132371.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 02-JUN-1999; 99WO-US12252.  
 PR 23-JUN-1999; 99US-0141037.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 28-JUL-1999; 99US-0146222.  
 PR 01-SEP-1999; 99WO-US20111.  
 PR 08-SEP-1999; 99WO-US20594.  
 PR 13-SEP-1999; 99WO-US20944.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.

PR 05-OCT-1999; 99WO-US23089.  
 PR 29-OCT-1999; 99US-0162506.  
 PR 29-NOV-1999; 99WO-US28214.  
 PR 30-NOV-1999; 99WO-US28313.  
 PR 30-NOV-1999; 99WO-US28409.  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31274.  
 PR 05-JAN-2000; 2000WO-US00219.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;  
 PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;  
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;  
 XX  
 DR WPI; 2000-572271/53.  
 DR N-PSDB; AAC58583.  
 XX  
 PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of  
 PT immune related disorders, e.g. systemic lupus erythematosus, Rheumatoid  
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -  
 XX  
 PS Claim 33; Fig 10; 309pp; English.  
 XX  
 CC The present invention describes sixty four human PRO proteins which can  
 CC be used in the treatment of immune related diseases. The human PRO  
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for  
 CC treating and diagnosing immune related disorders. The disorders are  
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,  
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,  
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's  
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic  
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,  
 CC immune-mediated renal disease, demyelinating diseases of the central  
 CC and peripheral nervous systems, hepatobiliary diseases, inflammatory  
 CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,  
 CC autoimmune or immune-mediated skin diseases, allergic diseases,  
 CC immunological diseases of the lung, and transplantation associated  
 CC diseases including graft rejection and graft-versus-host-disease.  
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used  
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and  
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein  
 CC sequences given in the exemplification of the present invention.  
 XX  
 SQ Sequence 443 AA;

Query Match 99.7%; Score 2505; DB 21; Length 443;  
 Best Local Similarity 99.8%; Pred. No. 5.6e-135;  
 Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLPASCCLPGSLLLWALLLLGLSASPQDSEEPDSTECTDGYEWDPSQHCQRDVNECLT 60  
 |||||  
 Db 1 mlpcasclpgsllllwallllllglssaspqdsdpstectdgyewdpdqshcdrvneclt 60  
 |||||  
 QY 61 IPEACKGEMKCNHYGYLCLPRSAAVINDLHGEGPPPPVPPAQHPNCPGPGYEDDQDS 120  
 |||||  
 Db 61 ipeackgemkcinygylclprsaavindlhgeggppppvppaghpnpccpgyepddqds 120  
 |||||  
 QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRQCRCVNL 180  
 |||||

Db 121 cvdvdecaqalhdcrpsqdcnlpssyqctcpdgyrkigpecvdiidecryrycqhrvcnl 180  
QY 181 PGFRCQCEPGFQLGPNNRSCVDNCEDMGAPCEQRCFNSYGTFLCRCHQGVELHRDGFs 240  
Db 181 pgsfrqccepgfqlgpnrrscvndnecdmgapceqrcfnsygtflcrchqgyvelhrdgs 240  
QY 241 CSDIDECSSYSLCQYRCVNEPGRFSCHCPQGYQLLATRLCQDIDCEGSAHQCSAQTC 300  
Db 241 csdidecsysylcgyrcvnepgfrfscpcpgyqllatrlcqdidecesgahqcseaqtc 300  
QY 301 VNFHGGYRCVDTNRCVEPYIQVSENRCCLCPASNPCLCREQSPSSIVHRYMTITSERSVPADV 360  
Db 301 vnfhggycvdtncrcvepyiqvsenrcclcpasnplcreqspssivhrymtitsersvpadv 360  
QY 361 FQIQATSVYPGAYNAFQIRAGNSQGDYFIQINNVFAMLVLPVTPGPREYVLDLEWVTM 420  
Db 361 fqiqatsvypgaynafqiragnsggdgyfirqinnvsamlvlpvtpgpreyvidlemvmt 420  
QY 421 NSLMSYRASSVLRILTVPFGAYTF 443  
Db 421 nslmsyrassvrlriltvfvgaytf 443

RESULT 4  
AA184707  
ID AAY84707 standard; Protein; 443 AA.  
XX AC AAY84707;  
XX XX  
DT 08-AUG-2000 (first entry)  
XX DE A human p53 mutant binding protein 1 (MBP1).  
XX KW Oncogene; p53; cell growth; immortalizing oncoprotein; cancer;  
KW p53 mutation; cell cycle; p53 mutant binding protein 1; MBP1.  
XX OS Homo sapiens.  
XX PN W0200022120-A1.  
XX PD 20-APR-2000.  
XX PF 12-OCT-1999; 99WO-FR02465.  
XX PR 12-OCT-1998; 98FR-0012754.  
XX PR 03-MAY-1999; 99US-0132331.  
XX PA (AVET ) AVENTIS PHARMA SA.  
XX PI Conseiller E, Debussche L, Gallagher W;  
XX DR WPI; 2000-329164/28.  
XX DR N-PSDB; AAA14608.  
XX PT New polypeptide that interacts with mutant p53, useful for developing  
PT anticancer agents such as antibodies, also stimulates cell growth and  
PT inhibits wild-type p53  
XX PS Claim 3; Page 84-85; 103pp; French.  
XX SS  
XX CC The specification describes a polypeptide that interacts specifically  
CC with oncogenic forms of p53, stimulates cell growth and blocks the  
CC antiproliferative action of wild-type p53. The polypeptide functions  
CC as an immortalizing oncoprotein and has synergistic interaction with  
CC oncogenic forms of p53 (both as regards oncogenic potential and  
CC proliferative effects). It is thus a target for treatment of cancers,  
CC particularly those associated with p53 mutations. The polypeptide is  
CC used to raise specific antibodies, and to screen for (or in rational  
CC design of) agents that modulate its interaction with oncogenic forms  
CC of p53. The antibodies and these agents are useful for treating diseases  
CC that involve abnormal functioning of the cell cycle, specifically cancer.  
CC The present sequence represents a p53 mutant binding protein 1 (MBP1),  
CC which is a protein of the invention.

XX SQ Sequence 443 AA;  
Query Match 99.7%; Score 2505; DB 21; Length 443;  
Best Local Similarity 99.8%; Pred. No. 5.6e-135;  
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLPASCCLPGSLLWALLLLLSASPOSEFPDSEPTCTDGYEWDPSQHCARDVNECLT 60  
Db 1 mlpcascclpgsllwlllllllgsaspqseepdsyctectdgyewdpdqshcdrvneclt 60  
QY 61 IPEACKGEMKCNHGYGCLPRSAVINDLHGEGLPPPPVPPAQHPNCPGPEYDQDS 120  
Db 61 ipeackgemkcnhgygylclprsaavindlhgeglppppvppaqhpnpcppgyepdqds 120  
QY 121 CVDVDECAQALHDCRPSQDCNLPSSYQCTCPDGYRKIGPECVDIIDECRYRYCQHRVCNL 180  
Db 121 cvdvdecaqalhdcrpsqdcnlpssyqctcpdgyrkigpecvdiidecryrycqhrvcnl 180  
QY 181 PGFRCQCEPGFQLGPNNRSCVDNCEDMGAPCEQRCFNSYGTFLCRCHQGVELHRDGFs 240  
Db 181 pgsfrqccepgfqlgpnrrscvndnecdmgapceqrcfnsygtflcrchqgyvelhrdgs 240  
QY 241 CSDIDECSSYSLCQYRCVNEPGRFSCHCPQGYQLLATRLCQDIDCEGSAHQCSAQTC 300  
Db 241 csdidecsysylcgyrcvnepgfrfscpcpgyqllatrlcqdidecesgahqcseaqtc 300  
QY 301 VNFHGGYRCVDTNRCVEPYIQVSENRCCLCPASNPCLCREQSPSSIVHRYMTITSERSVPADV 360  
Db 301 vnfhggycvdtncrcvepyiqvsenrcclcpasnplcreqspssivhrymtitsersvpadv 360  
QY 361 FQIQATSVYPGAYNAFQIRAGNSQGDYFIQINNVFAMLVLPVTPGPREYVLDLEWVTM 420  
Db 361 fqiqatsvypgaynafqiragnsggdgyfirqinnvsamlvlpvtpgpreyvidlemvmt 420  
QY 421 NSLMSYRASSVLRILTVPFGAYTF 443  
Db 421 nslmsyrassvrlriltvfvgaytf 443

RESULT 5  
AAY55850  
ID AAY55850 standard; protein; 443 AA.  
XX AC AAY55850;  
XX DT 13-MAR-2000 (first entry)  
XX DE Human SI-5 ECOMP-like protein (SELP).  
XX KW SI-5 extracellular matrix protein; SI-5 ECOMP-like protein; SELP; ECOMP;  
KW acquired immune deficiency syndrome; AIDS; allergy; asthma; human;  
KW irritable bowel syndrome; multiple sclerosis; osteoporosis; thyroiditis;  
KW cancer; infection; neoplastic disorder; immunological disorder.  
XX OS Homo sapiens.  
XX PN US6004753-A.  
XX PD 21-DEC-1999.  
XX PF 01-DEC-1997; 97US-0980514.  
XX PR 01-DEC-1997; 97US-0980514.  
XX PA (INCY-) INCYTE PHARM INC.  
XX PI Yue H, Shah P, Guegler KJ;  
XX DR WPI; 2000-072060/06.  
XX DR N-PSDB; AAZ39800.

PT Isolated polynucleotide and encoded polypeptide, useful for diagnosis  
PT and treatment of immunological and/or neoplastic disorders -  
XX  
XX  
XX Disclosure; Fig 1A-D; 29pp; English.  
XX  
XX This represents a human SL-5 extracellular matrix protein (ECMP)-like  
CC protein (SELP). The SELP DNA is useful for diagnosis of disorders  
CC associated with expression of SELP, e.g. acquired immune deficiency  
CC syndrome (AIDS), allergies, asthma, irritable bowel syndrome, multiple  
CC sclerosis, osteoporosis, thyroiditis, complications of cancer, viral,  
CC fungal, bacterial and protozoal infections, and neoplastic disorders such  
CC as adenocarcinoma, leukemia, lymphoma, and melanoma. The SELP DNA may  
CC also be used for Northern or Southern analysis, dot blot, or other  
CC membrane based technologies, or in dipstick, pin, ELISA (enzyme linked  
CC immunosorbent assay) or microarrays utilizing fluid or tissue from  
CC patient biopsies to detect altered expression of the protein. Antagonists  
CC of the SELP may be administered to a subject to treat or prevent an  
CC immunological and or neoplastic disorder.  
XX  
XX Sequence 443 AA;  
SQ  
Query Match 99.7%; Score 2505; DB 21; Length 443;  
Best Local Similarity 99.8%; Pred. No. 5.6e-135;  
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLPACSLPGSLLLWALLLLGASPDSEPDSTTECTDGYEWDPSQHCRRVNECLT 60  
Db 1 mlpcasclpgslllwlllllgaspsdsepdsttectdgyewdpdqhcrrvneclt 60  
QY 61 IPACKGEMKCNHYGYICLPESAVIDLHGEPPPPVPPPAOHNPCCPPGVEPDQDS 120  
Db 61 ipeackgemkcnhygyiclpesaavidlhgeppppvppppahnpccppgvepdqds 120  
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCQHCRCVNL 180  
Db 121 cvdvdecaqalhdcrpsqdchnlpgsyqctcpdgyrkigpecvdiidecryrycqhcrcvnl 180  
QY 181 PGSFRCQCEPGFOLGNNSCDVNECDMGAPCEQRCFNSYGFILCRHGCHYELHRDGF 240  
Db 181 pgsfrcqcepgfqlgnnscdvnecdmgapceqrcfnsygtflcrhgchgyelhrdgrfs 240  
QY 241 CSDIDECSSYSLCYRCVNEPGRFSCPCQGYOLLATRLCQDIDECESGAHCSEAQTC 300  
Db 241 csdidecssysylcyrcvnepgrfscpcqgyollatrlcqdidecesgahcseaqt 300  
QY 301 VNFHGYRCVDTNRCVEPYIQVSENRLCPASNPLCREQSPSSIVHRYMTITRSRVPADV 360  
Db 301 vnfhgyrcvdtncrvepyiqvsenrlcpasnplcreqspssivhrymtitrsersvpadv 360  
QY 361 FQIQATSVPGAYNAFQIRAGNSQGFYIRQINNFMVLAFLARPVTGPREYVLDLEWVTM 420  
Db 361 fqiqatsvypgaynafqiragnsqgfyirqinnvsamlvlarptvgpreyvidlenvmt 420  
QY 421 NSLMSYRASSVLRLLTVFGAYTF 443  
Db 421 .nslmsyrassvrlrtvfgaytf 443  
RESULT 6  
AAU12330  
ID AAU12330 standard; Protein; 443 AA.  
XX  
XX AC AAU12330;  
XX  
XX 24-OCT-2001 (first entry)  
XX  
XX DE Human PR0226 polypeptide sequence.  
XX  
XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
KW adipocyte; A-peptide; factor VIIA; gene therapy.

XX Homo sapiens.  
XX WO200140466-A2.  
XX 07-JUN-2001.  
XX 01-DEC-2000; 2000WO-US32678.  
XX 01-DEC-1999; 99WO-US28301.  
XX 01-DEC-1999; 99WO-US28634.  
XX 02-DEC-1999; 99WO-US28551.  
XX 02-DEC-1999; 99WO-US28564.  
XX 02-DEC-1999; 99WO-US28565.  
XX 09-DEC-1999; 99US-0170262.  
XX 16-DEC-1999; 99WO-US30095.  
XX 20-DEC-1999; 99WO-US30911.  
XX 20-DEC-1999; 99WO-US30999.  
XX 30-DEC-1999; 99WO-US31243.  
XX 06-JAN-2000; 2000WO-US00277.  
XX 06-JAN-2000; 2000WO-US00376.  
XX 11-FEB-2000; 2000WO-US03565.  
XX 18-FEB-2000; 2000WO-US04341.  
XX 18-FEB-2000; 2000WO-US04342.  
XX 22-FEB-2000; 2000WO-US04414.  
XX 24-FEB-2000; 2000WO-US04914.  
XX 01-MAR-2000; 2000WO-US05601.  
XX 20-MAR-2000; 2000WO-US07377.  
XX 21-MAR-2000; 2000WO-US07532.  
XX 30-MAR-2000; 2000WO-US08439.  
XX 17-MAY-2000; 2000WO-US13705.  
XX 22-MAY-2000; 2000WO-US14042.  
XX 30-MAY-2000; 2000WO-US14941.  
XX 02-JUN-2000; 2000WO-US15264.  
XX 10-NOV-2000; 2000WO-US30873.  
XX (GETH ) GENENTECH INC.  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
WPI: 2001-408281/43.  
N-PSDB; AAS21402.  
XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
PT other PRO polypeptides, link bioactive molecules to cells expressing  
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
PT lung, breast, prostate, cervical -  
XX Claim 12; Fig 318; 813pp; English.  
XX AAU12172-AAU12446 represent novel human secretory and transmembrane  
CC PRO polypeptides. The PRO polypeptides are useful to detect other  
CC PRO polypeptides, to link bioactive molecules to cells expressing  
CC PRO polypeptides, to modulate biological activities of cells expressing  
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
CC polypeptide expression in a cell sample to that in a control sample.  
CC Some of the 275 sequences are also useful to stimulate the release of  
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
CC proliferation or differentiation of chondrocytes, the proliferation or  
CC gene expression in pericyte cells, the release of proteoglycans from  
CC cartilage, the proliferation of inner ear utricular supporting cells or  
CC of T-lymphocytes, the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.

```
XX SQ Sequence 443 AA;
Query Match 99.7%; Score 2505; DB 22; Length 443;
Best Local Similarity 99.8%; Pred. No. 5,6e-135;
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLPASCCLPGSLLLWALLLLLLGSSASQDSEPDSTCTDGYEWDPSOHCRCVNECLT 60
DB 1 mlpcascclpgsllllwalllllllgssasqdsdpstctdgyewdpshqcrdvneclt 60
QY 61 IPEACKGEMKCNHGYGCLPRSAVINDLHGEGPPPPVPPAQAHPNCPGPGYEPDQDS 120
DB 61 ipeackgemkcinhygyiclprsaavindlhgegppppvppaqaahpncppgypdpdqs 120
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRCQHRVNL 180
DB 121 cvdvdecaqalhdcrpsqdchnlpgsyqctcpdgyrkigpecvdiidecryrcqhrvnl 180
QY 181 PGSFRCQCEPGFOLGPNNSRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGS 240
DB 181 pgsfrqccepgfolgpnnsrscvdvnecdmgapceqrcfnsygtflcrchqgyelhrdgs 240
QY 241 CSDIDECSSYSSYLCOYRCVNEPGRFSCHPQGYQLLALRLCQDIDECESGAHQCSAQTC 300
DB 241 csdidecsyssylcoyrcvnepgrfscchpqqgyqlalrlcqdidecesgahqcseaqt 300
QY 301 VNFHGGYRCVDTNRCVPEYIQVSENRCPLCPASNPLOREOPSSIVHRYMTITSSERSVPADV 360
DB 301 vnfhgyrcvdtncrcvepyiqvsenrcplcpasnplcreqpsivhrymtitssersvpadv 360
QY 361 FOIQATSVYPGAYNAFQIRAGNSQGDYFIQINNVFAMLVLARPVTPGREYVLDLEMTM 420
DB 361 fqiqtatvypgaynafqiragnsqgdyfiqrinnvsamlvlarpvtpgreyvldlemvtm 420
QY 421 NSLMSYRASSVLRITVFGAYTF 443
DB 421 nslmsyrassvrlrtvfgaytf 443

RESULT 7
AAB92533
ID AAB92533 standard; Protein; 443 AA.
XX AC AAB92533;
XX DT 26-JUN-2001 (first entry)
XX DE Human protein sequence SEQ ID NO:10694.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX OS Homo sapiens.
XX PN EF1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-0116126.
XX PR 29-JUL-1999; 99JP-0248036.
XX PR 27-AUG-1999; 99JP-0300253.
XX PR 11-JAN-2000; 2000JP-0118776.
XX PR 02-MAY-2000; 2000JP-0183767.
XX PR 09-JUN-2000; 2000JP-0241899.
XX PR (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
```

```
XX PT primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX Claim 8; SEQ ID 10694; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
```

SQ Sequence 443 AA;

Query Match 99.3%; Score 2495; DB 22; Length 443;  
Best Local Similarity 99.5%; Pred. No. 2.1e-134;  
Matches 441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MLPASCCLPGSLLLWALLLLLLGSSASQDSEPDSTCTDGYEWDPSOHCRCVNECLT 60
DB 1 mlpcascclpgsllllwalllllllgssasqdsdpstctdgyewdpshqcrdvneclt 60
QY 61 IPEACKGEMKCNHGYGCLPRSAVINDLHGEGPPPPVPPAQAHPNCPGPGYEPDQDS 120
DB 61 ipeackgemkcinhygyiclprsaavindlhgegppppvppaqaahpncppgypdpdqs 120
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRCQHRVNL 180
DB 121 cvdvdecaqalhdcrpsqdchnlpgsyqctcpdgyrkigpecvdiidecryrcqhrvnl 180
QY 181 PGSFRCQCEPGFOLGPNNSRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGS 240
DB 181 pgsfrqccepgfolgpnnsrscvdvnecdmgapceqrcfnsygtflcrchqgyelhrdgs 240
QY 241 CSDIDECSSYSSYLCOYRCVNEPGRFSCHPQGYQLLALRLCQDIDECESGAHQCSAQTC 300
DB 241 csdidecsyssylcoyrcvnepgrfscchpqqgyqlalrlcqdidecesgahqcseaqt 300
QY 301 VNFHGGYRCVDTNRCVPEYIQVSENRCPLCPASNPLOREOPSSIVHRYMTITSSERSVPADV 360
DB 301 vnfhgyrcvdtncrcvepyiqvsenrcplcpasnplcreqpsivhrymtitssersvpadv 360
QY 361 FOIQATSVYPGAYNAFQIRAGNSQGDYFIQINNVFAMLVLARPVTPGREYVLDLEMTM 420
DB 361 fqiqtatvypgaynafqiragnsqgdyfiqrinnvsamlvlarpvtpgreyvldlemvtm 420
QY 421 NSLMSYRASSVLRITVFGAYTF 443
DB 421 nslmsyrassvrlrtvfgaytf 443
```

RESULT 8  
AAY84706

ID	AA584706 standard; Protein; 443 AA.	Db	241	csdidecgyssylcgyrcvneprfscchpcpgyllatrlcqdideceghagcseaqt	300
XX	AA584706;	QY	301	VNFHGGYRCVDTNRCVEPIQVSENRCCLCPASNPCLREQPSSIVHRYMTITSERSVPADV	360
XX	08-AUG-2000 (first entry)	Db	301	vnfhgyrcvdtncrcvpepyqvsdnrcclcpasnpclreqpssivhrymsitsersvpadv	360
DE	Amino acid sequence of a p53 mutant binding protein 1 (MBP1).	QY	361	FQIQATSVTPGAYNARQIRAGNSQGFYIRQINNVFAMLVLARPVTPGPREYVLDLEKVTM	420
XX	Oncogene; p53; cell growth; immortalizing oncoprotein; cancer;	Db	361	fqiqtavpypgaynafqirsgntqgdfyirqinnvsamvlarpvtgpreyvidlemvmt	420
KW	p53 mutation; cell cycle; p53 mutant binding protein 1; MBP1.	QY	421	NSLMSYRASSVLRLTVFVGAYTF 443	
XX	Mus sp.	Db	421	nslmsyrassvrltvtfvgaytf 443	
XX	WO2000022120-A1.	RESULT	9		
XX	20-APR-2000.	AA58353			
XX	12-OCT-1999; 99WO-PR02465.	ID	AA58353 standard; Protein; 433 AA.		
XX	12-OCT-1998; 98PR-0012754.	XX			
PR	03-MAY-1999; 99US-0132331.	AC	AA58353;		
XX	(AVET ) AVENTIS PHARMA SA.	XX			
PA	Conseiller E, Debussche L, Gallagher W;	DT	14-MAR-2001 (first entry)		
XX		XX			
PI		DE	Lung cancer associated polypeptide sequence SEQ ID 691.		
XX		XX			
DR	WPI: 2000-329164/28.	XX			
DR	N-PSDB; AAA14603.	KW	Human; lung cancer associated protein; neuroprotective; cytostatic;		
XX		KW	cardioactive; immunomodulatory; muscular active; vulnerary;		
XX		KW	gastrointestinal; nephrotropic; antiinfective; gynecological;		
PT	New polypeptide that interacts with mutant p53, useful for developing	KW	antibacterial; diagnosis; neural disorder; immune disorder; reproductive;		
PT	anticancer agents such as antibodies, also stimulates cell growth and	KW	proliferative disorder; wound healing; infectious disease.		
PT	inhibits wild-type p53 -	XX			
XX		OS	Homo sapiens.		
PS	Claim 2; Page 79-80; 103pp; French.	XX			
XX		PN	WO200055180-A2.		
XX	The specification describes a polypeptide that interacts specifically	PD	21-SEP-2000.		
CC	with oncogenic forms of p53, stimulates cell growth and blocks the	XX			
CC	antiproliferative action of wild-type p53. The polypeptide functions	PD			
CC	as an immortalizing oncoprotein and has synergistic interaction with	PF	08-MAR-2000; 2000WO-US05918.		
CC	oncogenic forms of p53 (both as regards oncogenic potential and	XX			
CC	proliferative effects). It is thus a target for treatment of cancers,	PR	12-MAR-1999; 99US-0124270.		
CC	particularly those associated with p53 mutations. The polypeptide is	XX			
CC	used to raise specific antibodies, and to screen for (or in rational	XX	(HUMA-) HUMAN GENOME SCI INC.		
CC	design of) agents that modulate its interaction with oncogenic forms	PA	(ROSE/) ROSEN C A.		
CC	of p53. The antibodies and these agents are useful for treating diseases	PI	Ruben SM;		
CC	that involve abnormal functioning of the cell cycle, specifically cancer.	XX			
CC	The present sequence represents a p53 mutant binding protein 1 (MBP1),	XX			
CC	which is a protein of the invention.	DR	WPI: 2000-587514/55.		
XX		DR	N-PSDB; AAF18229.		
SQ	Sequence 443 AA;	XX			
		PT	Lung cancer associated gene sequences, referred to as lung cancer		
		PT	antigens, useful for treatment, prevention, and diagnosis of disorders		
		PT	such as lung cancer -		
		XX			
		PS	Claim 11; Page 1197-1199; 1425pp; English.		
		XX			
QY	1 MIPACSLPGSLLWALLLLGASPDSEPDSTVCTDGYEWDPSQHCVRVNECLT 60	CC	Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer		
Db	1 mlpfascplpgsllwaflllllgaaspqdpdpdytctdgyewdadsghcrdvneclt 60	CC	associated proteins represented in AAB58106 - AAB58548. Lung cancer		
QY	61 IPEACKGEMKCNHYGYLCLPSAAVINDLHCEGPPPPVPPAQHPNCPGVEPDDQDS 120	CC	associated proteins and polynucleotide sequences, their agonists, and		
Db	61 ipeackgemkcnhygyiclprsaavisdllhgeggppppaahqpnpcqgpepdeqs 120	CC	antagonists may have neuroprotective; cytostatic; cardioactive;		
QY	121 CVDVDECAALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYCOHRCVNL 180	CC	immunomodulatory; muscular active general; vulnerary; gastrointestinal		
Db	121 cvdvdectalhdcrpsqdchmlpgsyqctcdpgyrkigpecvdiidecryyqcrcvnl 180	CC	general; nephrotropic; antiinfective; gynecological; or antibacterial		
QY	181 PGSFRCQCEPFGQLGNPNRSCVDVNECDMGAPCEQRCFNSYGTFLCRHQGYELHRDGF 240	CC	activity. The invention also includes antibodies specific for the		
Db	181 pgsfrcqcepgfqlgnpnrscvdvnecdmgapceqrccfnsgygtflcrnqgyelhrdgs 240	CC	protein or polynucleotide sequences. The lung cancer associated		
QY	241 CSDIDECSTSSYLQYRCVNEPRGFSCHCPQGYQLLATRLCQDIDECESGAHQCSAQTC 300	CC	polynucleotide sequences may be used for detection of lung cancer,		
		CC	chromosome identification, as chromosome markers, and for numerous other		
		CC	diagnostic or research purposes. The proteins may be used to treat		
		CC	disorders such as neural, immune, muscular, reproductive,		
		CC	gastrointestinal, pulmonary, cardiovascular, renal, and proliferative		
		CC	disorders. The proteins may also be used in the treatment of wounds and		
		CC	infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and		
		CC	peptide AAB58549 are used in the course of the invention for the		
		CC	identification and characterisation of the polynucleotide and protein		





DR N-PSDB; AAL14601.  
 XX  
 PT New polypeptide that interacts with mutant p53, useful for developing  
 PT anticancer agents such as antibodies, also stimulates cell growth and  
 PT inhibits wild-type p53 -  
 XX  
 PS Claim 20; Page 74-75; 103pp; French.  
 XX  
 CC The specification describes a polypeptide that interacts specifically  
 CC with oncogenic forms of p53, stimulates cell growth and blocks the  
 CC antiproliferative action of wild-type p53. The polypeptide functions  
 CC as an immortalizing oncoprotein and has synergistic interaction with  
 CC oncogenic forms of p53 (both as regards oncogenic potential and  
 CC proliferative effects). It is thus a target for treatment of cancers,  
 CC particularly those associated with p53 mutations. The polypeptide is  
 CC used to raise specific antibodies, and to screen for (or in rational  
 CC design of) agents that modulate its interaction with oncogenic forms  
 CC of p53. The antibodies and these agents are useful for treating diseases  
 CC that involve abnormal functioning of the cell cycle, specifically cancer.  
 CC The present sequence represents the C-terminal of a p53 mutant binding  
 CC protein 1 (C-mbpl), which is a protein of the invention.  
 XX  
 SQ Sequence 295 AA;

Query Match 64.4%; Score 1619; DB 21; Length 295;  
 Best Local Similarity 96.9%; Pred. No. 7.8e-85;  
 Matches 286; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 149 CTCPDGYRIGPCVDPIDECRYCOHRCVNLPGSFRCCQCEPGFOLGPNRSCVDVNECD 208  
 Db 1 ctcpgdyrkigpcvdpidecryrcvncvnlpgsfrccqcepgfqlgpnrscvndvnecd 60

QY 209 MGAPCEQRCFNSYGTFLCRCHOGYELHRDGFSCSDIDECYSYLCQYRCVNEPGRFSCH 268  
 Db 61 mgapceqrcfnsygtflcrchogylhrdghfscsdidecysylcqyrcvnepggrfsc 120

QY 269 CPQGYOLLATRLCQDIDECESGAHQCEAQTCVNFHGGYRCVDTNRCVPEYIQVSENRLC 328  
 Db 121 cpqgyollatrlcqdidecetgahqceaqtcvnfhggycvdtncrcvpeyqvgsdnrlc 180

QY 329 CPASNPLCREQPSIVHRYMTITTSERSVPADVFQIQATSVYPGAYNAFQIRAGNSQGDYF 388  
 Db 181 cpasnplcreqpsivhrymtsersvpadvfqiqtatsvypgaynafqirsgntgdyf 240

QY 389 IROINNVMFAMVLARVTPGREYVLDLEMTNMSLSYRASSVRLRTVFGAYTF 443  
 Db 241 irqinnvsmvlarpvtgpreyvldlemtnmslsyrassvrlrtvfgaytf 295

RESULT 12  
 AAY76081  
 ID AAY76081 standard; Protein; 274 AA.  
 XX  
 AC AAY76081;  
 XX  
 DT 27-MAR-2000 (first entry)  
 XX  
 DE Human EGF extracellular homologue, SEQ ID NO:336.  
 XX  
 KW Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;  
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;  
 KW secreted; transmembrane; inflammation; cancer; neurological disease;  
 KW angiogenesis; tumour vascularisation; growth disorder;  
 KW developmental disorder; skin wound; hair follicle disorder;  
 KW anti-inflammatory; cytostatic; neuroprotective; vulnery.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9955865-A1.  
 XX  
 PD 04-NOV-1999.  
 XX

PF 29-APR-1999; 99WO-NZ00051.  
 XX  
 PR 29-APR-1998; 98US-0069726.  
 PR 09-NOV-1998; 98US-0188930.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 PI Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;  
 DR WPI; 2000-072177/06.  
 DR N-PSDB; AA61786.  
 XX  
 PT Novel polynucleotides useful for the treatment of various conditions  
 PT including wounds and cancer -  
 XX  
 PS Claim 4; Page 197-198; 235pp; English.  
 XX  
 CC The invention relates to novel nucleic acid sequences derived from rat  
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,  
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying  
 CC cells. Polypeptides of the invention may be used to treat inflammation,  
 CC cancer and neurological diseases. The proteins may be used to stimulate  
 CC the growth and motility of keratinocytes, to inhibit the growth of  
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to  
 CC modulate skin inflammation, to modulate epithelial cell growth and to  
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used  
 CC to treat growth and developmental defects, skin wounds and hair follicle  
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded  
 CC by cDNA sequences derived from several mouse, rat or human skin cell  
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and  
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating  
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,  
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more  
 CC putative transmembrane domains.  
 XX  
 SQ Sequence 274 AA;

Query Match 60.4%; Score 1518; DB 21; Length 274;  
 Best Local Similarity 99.6%; Pred. No. 4e-79;  
 Matches 273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 YRYCQHRVCNVLPGSFRCCQCEPGFOLGPNRSCVDVNECDMGAPCEQRCFNSYGTFLCRCH 229  
 Db 1 yrycqhrvcnvlpgsfrccqcepgfqlgpnrscvndvnecdmgapceqrcfnsygtflcrch 60

QY 230 QGYELHRDGFSCSDIDECYSYLCQYRCVNEPGRFSCHPCQGYOLLATRLCQDIDECES 289  
 Db 61 qgyelhrdghfscsdidecysylcqyrcvnepggrfscpcqgyollatrlcqdideces 120

QY 290 GAHQCEAQTCVNFHGGYRCVDTNRCVPEYIQVSENRLCPASNPLCREQPSIVHRYMT 349  
 Db 121 gahqceaqtcvnfhggycvdtncrcvpeyiqvsenrlcpasnplcreqpsivhrymt 180

QY 350 ITTSERSVPADVFQIQATSVYPGAYNAFQIRAGNSQGDYFIRQINNVMFAMVLARVTPGR 409  
 Db 181 ittsersvpadvfqiqtatsvypgaynafqiragsgdgyirqinnvsmvlarpvtgpr 240

QY 410 EYVLDLEMTNMSLSYRASSVRLRTVFGAYTF 443  
 Db 241 eyvldlemtnmslsyrassvrlrtvfgaytf 274

RESULT 13  
 AAB56020  
 ID AAB56020 standard; Protein; 274 AA.  
 XX  
 AC AAB56020;  
 XX  
 DT 08-MAR-2001 (first entry)  
 XX  
 DE Skin cell protein, SEQ ID NO: 336.  
 XX



Db 424 ntfrksgnengefyrqtspvsmvlvkslsgprehivdmltvtssigtftssvrlr 483

QY 434 LTVFVGAYTF 443

Db 484 ltiivgpfsf 493

II: II ::I

RESULT 15

AAB48077

ID AAB48077 standard; protein; 493 AA.

XX AAB48077;

AC

DT 19-MAR-2001 (first entry)

XX

DE Human extracellular signaling molecule (EXCS) (ID 1359783CD1).

XX

KW Extracellular signaling molecule; EXCS; anti-inflammatory; human; immunosuppressive; cytostatic; neuroprotective; gastrointestinal; viricide; antibacterial; anti-HIV; human immunodeficiency virus; antifertility; cerebroprotective; nootropic; antiulcer; antifungal; anticonvulsant; tranquilizer; neuroleptic; vasotropic; gynecological; keratolytic; protozoacide; gene therapy.

XX

OS Homo sapiens.

XX

PN WO200070049-A2.

XX

PD 23-NOV-2000.

XX

XX 19-MAY-2000; 2000WO-US13975.

PR 15-JUL-1999; 99US-0134949.

PR 30-JUL-1999; 99US-0144270.

PR 04-OCT-1999; 99US-0146700.

XX 99US-0157508.

XX (INCY-) INCYTE GENOMICS INC.

XX

PI Tang YT, Yue H, Lal P, Burford N, Bandman O, Baughn MR;

PI Azlmzai Y, Lu DAM, Patterson C;

XX

DR WPI: 2001-025021/03.

DR N-PSDB; AAC84313.

XX

PT New human extracellular signaling nucleic acids and polypeptides useful for diagnosing, treating and preventing infections and gastrointestinal, neurological, reproductive, and autoimmune/inflammatory disorders -

XX

PS Claim 1; Page 96-97; 114pp; English.

XX

CC The invention provides human extracellular signaling molecules (EXCS) and polynucleotides which identify and encode EXCS. EXCS can be expressed by standard recombinant methodology. The amino acid and nucleic acid sequences of EXCS are useful for diagnosing, treating and preventing infections and gastrointestinal (peptic ulcer, dysphagia, pancreatitis), neurological (e.g. epilepsy, ischemic cerebrovascular disease, stroke), reproductive (infertility, ovulatory defects, endometriosis), autoimmune/inflammatory (actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease), and cell proliferative disorders including cancers (of the breast, adrenal gland, bone). They may also be used to treat fatal familial insomnia, nutritional and metabolic diseases of the nervous system, myopathies, mental disorders (anxiety, schizophrenia, mood), as well as infections caused by parasites (malaria, leishmania, trypanosoma), viral (adenovirus, coronavirus, flavivirus), bacterial (e.g. pneumococcus, staphylococcus, bacillus), and fungal (aspergillus, blastomycosis, dermatophytes) agents. The nucleic acids, polypeptides, antagonists, agonists, pharmaceutical compositions, and antibodies may also be used for treating or preventing disorders associated with increased or decreased expression or activity of EXCS. EXCS polynucleotides may also be used to detect and quantify gene expression in biopsied tissues in

CC which expression of EXCS may be correlated with the disease, to determine presence or excess expression of EXCS, to monitor regulation of EXCS levels during therapeutic intervention, to detect the presence of CC associated disorders, as targets in microarray, to generate hybridization probes, and to detect differences in gene sequences among normal, carrier or affected individuals. Antibodies may also be used in diagnosing CC disorders, in monitoring patients being treated with EXCS agonists, CC antagonists or inhibitors. Sequences AAB48057-B48082 represent the EXCS of the invention.

XX

SQ Sequence 493 AA;

Query Match 54.9%; Score 1379; DB 22; Length 493;

Best Local Similarity 48.8%; Pred. No. 5,le-71;

Matches 239; Conservative 76; Mismatches 109; Indels 66; Gaps 4;

QY 17 LLLLLLGSASPDSEPDSTVCTGDEWDPSOHCRODNECLTIPACKGEMKCNHYG 76

Db 7 lmltialvksqteetitytqctdgyewdvprqgckddecdivpdcackgmkcvnhyg 66

QY 77 GYLCLPRSAVINDLHGEGPPPPVPPAQ----- 104

Db 67 gylclphtaql--vnnepqgetpaegtsgattgvvaassmatsgvlpgggfvasaaa 124

QY 105 -----HPNCPPGYEPDQDSCVDVDECAQALHDC 134

Db 125 vagpemqtgrnnfvirnpadqripnspshriqcaagyeqsehnvcqdidectagthnc 184

QY 135 RPSQDCHNLPGSYOCTCPDGYRKIGPECVDIDECRY-RVCOHRCVNLPGSFRCCQEPGFQ 193

Db 185 radqvcinlrgsfacqppgykrgqecvldictlppyqhgrcvnptpsfyqcqspggf 244

QY 194 LGPNRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGFSCSIDECSSSYL 253

Db 245 laannytvdinecdasncacqgcynilgscfcqngyelsdrincedidecrtssyl 304

QY 254 QYRCVNEPGRFSCPCQGYQLLATRLCQDIDECESGAHCQSEAQTCVNFPHGGRVDTN 313

Db 305 cdyqcvnepgkfscmcpggygvvrstcqdinecet-tnecredemcwnyhgfrcyprn 363

QY 314 RCVEPYIQVSENRCLCPASNPCLREQPPSSIVHRYMTTTSERSVPADVFQIOATVYPGAY 373

Db 364 pqdpyilltpenrcvcpvsnamcrelpqsvlykysirsdsvpsdfiqattiyanti 423

QY 374 NAFQIRAGNSQGDIFYIRQINNFMVLARVPTGPREYVLDLEMTNMSLYRASVSLR 433

Db 424 ntfriksgnengefyrqtspvsmvlvkslsgprehivdmltvtssigtftssvrlr 483

QY 434 LTVFVGAYTF 443

Db 484 ltiivgpfsf 493

Search completed: September 25, 2002, 09:42:13

Job time: 248 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2002, 09:38:20 ; Search time 15.92 Seconds  
(without alignments)  
679.682 Million cell updates/sec

Title: US-09-829-936A-22  
Perfect score: 2513  
Sequence: 1 MLPSCSLGSLLLWALLLL.....MSTRASSVLRLTVFGAYTF 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2505	99.7	443	2	US-08-833-963C-2
2	2505	99.7	443	3	US-08-980-514-1
3	1518	60.4	274	4	US-09-188-930-336
4	1283	51.1	448	2	US-08-884-072-1
5	1283	51.1	448	4	US-09-212-168-1
6	1147	45.6	387	2	US-08-884-072-5
7	1147	45.6	387	2	US-08-833-963C-9
8	1147	45.6	387	3	US-08-980-514-3
9	1147	45.6	387	4	US-09-212-168-5
10	1053.5	41.9	337	4	US-09-188-930-186
11	523.5	20.8	1833	3	US-08-479-722B-2
12	523.5	20.8	1833	5	PCT-US95-02251-18
13	480	19.1	1394	6	5171197-30
14	461	18.3	1253	3	US-08-479-722B-4
15	452	18.0	1251	5	PCT-US95-02251-3
16	452	18.0	1252	1	US-08-199-780-3
17	452	18.0	1252	2	US-08-316-650-3
18	436	17.3	638	2	US-08-897-443-1
19	413.5	16.5	956	2	US-08-897-443-3
20	393	15.6	2556	1	US-08-185-432-17
21	393	15.6	2556	1	US-08-083-590A-20
22	393	15.6	2556	3	US-08-532-384-20
23	388	15.4	886	3	US-09-110-116-3
24	388	15.4	2523	1	US-08-185-432-18
25	377.5	15.0	2471	1	US-08-185-432-16
26	377.5	15.0	2471	1	US-08-083-590A-19
27	377.5	15.0	2471	3	US-08-532-384-19

28	358.5	14.3	2703	1	US-08-185-432-19	Sequence 19, Appl
29	356.5	14.2	816	3	US-08-820-170A-37	Sequence 37, Appl
30	356.5	14.2	816	3	US-09-055-699-37	Sequence 37, Appl
31	356.5	14.2	816	4	US-09-273-565-37	Sequence 37, Appl
32	356.5	14.2	816	4	US-09-565-538-37	Sequence 37, Appl
33	344.5	13.7	652	2	US-08-751-305-2	Sequence 2, Appli
34	327.5	13.0	810	2	US-08-820-170A-34	Sequence 34, Appl
35	327.5	13.0	810	3	US-09-055-699-34	Sequence 34, Appl
36	327.5	13.0	810	4	US-09-273-565-34	Sequence 34, Appl
37	327.5	13.0	810	4	US-09-565-538-34	Sequence 34, Appl
38	324.5	12.9	1193	2	US-08-400-159-10	Sequence 10, Appl
39	324.5	12.9	1193	3	US-08-611-729A-10	Sequence 10, Appl
40	323	12.9	673	1	US-08-282-141-3	Sequence 3, Appli
41	323	12.9	673	1	US-08-435-434-1	Sequence 1, Appli
42	323	12.9	673	1	US-08-435-436-1	Sequence 1, Appli
43	323	12.9	673	2	US-08-438-863-1	Sequence 1, Appli
44	323	12.9	673	2	US-08-438-864-1	Sequence 1, Appli
45	323	12.9	673	3	US-08-438-862-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-833-963C-2  
; Sequence 2, Application US/08033963C  
; Patent No. 5916769  
; GENERAL INFORMATION:  
; APPLICANT: Olsen, et al.  
; TITLE OF INVENTION: Extracellular/Epidermal Growth Factor  
; TITLE OF INVENTION: HCABA58X  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Ave  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/833,963C  
; FILING DATE: 11-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US96/05033  
; FILING DATE: 10-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PE258  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-309-8504  
; TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 443 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-833-963C-2

Query Match 99.7%; Score 2505; DB 2; Length 443;  
Best Local Similarity 99.8%; Pred. No. 2e-196;  
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLPSCSLGSLLLWALLLLLGASPDSEPDSTCTDGYEWDPSQCHRDVNECLT 60  
|||||

Db 1 MLPASCCLPGSLLLWALLLLLLLGSASPDSEPTDSTECTDGYEWDPSQHCRDNECLT 60  
QY 61 IPACKGEMKCINHYGYLCLPRSAVINDLHCEGPPPPVPPAHPNPPCPGYPEDDQDS 120  
Db 61 IPACKGEMKCINHYGYLCLPRSAVINDLHCEGPPPPVPPAHPNPPCPGYPEDDQDS 120  
QY 121 CVDVDECAALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCOHRCVNL 180  
Db 121 CVDVDECAALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCOHRCVNL 180  
QY 181 PGSFRCOCPEFGFOLGPNNRSCVDVNECDMGAPCEQRCFNSTGYFLCRCHQGYELHRDGS 240  
Db 181 PGSFRCOCPEFGFOLGPNNRSCVDVNECDMGAPCEQRCFNSTGYFLCRCHQGYELHRDGS 240  
QY 241 CSDIDECSSYLLCOYRCVNEPGRFSCPCQGYQLLATRLCQDIDECESGAHOCSEAQC 300  
Db 241 CSDIDECSSYLLCOYRCVNEPGRFSCPCQGYQLLATRLCQDIDECESGAHOCSEAQC 300  
QY 301 VNFHGGYRCVDTNRCVEPYIQVSENRCCLCPASNPLCREQSPSSIVHRYMTTTSERSVPADV 360  
Db 301 VNFHGGYRCVDTNRCVEPYIQVSENRCCLCPASNPLCREQSPSSIVHRYMTTTSERSVPADV 360  
QY 361 FQIQATSVYPCAYNAFQIRAGNSQGDYIRQINNVSAMLVLARPVTPGREYVLDLEWVTM 420  
Db 361 FQIQATSVYPCAYNAFQIRAGNSQGDYIRQINNVSAMLVLARPVTPGREYVLDLEWVTM 420  
QY 421 NSLMSYRASSVLRLTVFVGAYTF 443  
Db 421 NSLMSYRASSVLRLTVFVGAYTF 443

## RESULT 2

US-08-980-514-1  
; Sequence 1, Application US/08980514  
; Patent No. 6004753  
; GENERAL INFORMATION:  
; APPLICANT: Yue, Henry  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: HUMAN S1-5-ECMP-LIKE PROT  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980,514  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0436 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 443 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; IMMEDIATE SOURCE:  
; LIBRARY: BRSTNOT13  
; CLONE: 2786449  
US-08-980-514-1

Query Match 99.7%; Score 2505; DB 3; Length 443;  
Best Local Similarity 99.8%; Pred. No. 2e-196;  
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLPASCCLPGSLLLWALLLLLLLGSASPDSEPTDSTECTDGYEWDPSQHCRDNECLT 50  
Db 1 MLPASCCLPGSLLLWALLLLLLLGSASPDSEPTDSTECTDGYEWDPSQHCRDNECLT 60  
QY 61 IPACKGEMKCINHYGYLCLPRSAVINDLHCEGPPPPVPPAHPNPPCPGYPEDDQDS 120  
Db 61 IPACKGEMKCINHYGYLCLPRSAVINDLHCEGPPPPVPPAHPNPPCPGYPEDDQDS 120  
QY 121 CVDVDECAALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCOHRCVNL 180  
Db 121 CVDVDECAALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCOHRCVNL 180  
QY 181 PGSFRCOCPEFGFOLGPNNRSCVDVNECDMGAPCEQRCFNSTGYFLCRCHQGYELHRDGS 240  
Db 181 PGSFRCOCPEFGFOLGPNNRSCVDVNECDMGAPCEQRCFNSTGYFLCRCHQGYELHRDGS 240  
QY 241 CSDIDECSSYLLCOYRCVNEPGRFSCPCQGYQLLATRLCQDIDECESGAHOCSEAQC 300  
Db 241 CSDIDECSSYLLCOYRCVNEPGRFSCPCQGYQLLATRLCQDIDECESGAHOCSEAQC 300  
QY 301 VNFHGGYRCVDTNRCVEPYIQVSENRCCLCPASNPLCREQSPSSIVHRYMTTTSERSVPADV 360  
Db 301 VNFHGGYRCVDTNRCVEPYIQVSENRCCLCPASNPLCREQSPSSIVHRYMTTTSERSVPADV 360  
QY 361 FQIQATSVYPCAYNAFQIRAGNSQGDYIRQINNVSAMLVLARPVTPGREYVLDLEWVTM 420  
Db 361 FQIQATSVYPCAYNAFQIRAGNSQGDYIRQINNVSAMLVLARPVTPGREYVLDLEWVTM 420  
QY 421 NSLMSYRASSVLRLTVFVGAYTF 443  
Db 421 NSLMSYRASSVLRLTVFVGAYTF 443

## RESULT 3

US-09-188-930-336  
; Sequence 336, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods For Their Use  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 336  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Human  
US-09-188-930-336

Query Match 60.4%; Score 1518; DB 4; Length 274;  
Best Local Similarity 99.6%; Pred. No. 2.6e-116;  
Matches 273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 YRYCOHRCVNLPGSFRCQCEPGLGPNNRSCVDVNECDMGAPCEQRCFNSTGYFLCRCH 229  
|||||

Db 1 YRYCQHRVNLPGSFRCOCEPGFQGLGNRRSCVDYNECDMGAPCEQRCFNSYGTFLCRCH 60  
QY 230 QYELHROGFSDDIDECYSYLYCQYRCVNEPGRFSCHPGQYQALLATRLCQDIDECES 289  
Db 61 QYELHROGFSDDIDECYSYLYCQYRCVNEPGRFSCHPGQYQALLATRLCQDIDECES 120  
QY 290 GAHQCEAQTQVNFHGGYRCVDTNRCVEPIQVSENRCCLCPASNPLCRQPSIVHRYMT 349  
Db 121 GAHQCEAQTQVNFHGGYRCVDTNRCVEPIQVSENRCCLCPASNPLCRQPSIVHRYMT 180  
QY 350 ITSERSVPADVFOIQATSYPGAYNAFOIRAGNSQGDYFIROINNVMFAMLVLRPVGTGR 409  
Db 181 ITSERSVPADVFOIQATSYPGAYNAFOIRAGNSQGDYFIROINNVMFAMLVLRPVGTGR 240  
QY 410 EYVLDLMTMNSLMSYRASSVLRITVFGAYTF 443  
Db 241 EYVLDLMTMNSLMSYRASSVLRITVFGAYTF 274

## RESULT 4

US-08-884-072-1

; Sequence 1, Application US/08884072

; Patent No. 5872234

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Corley, Neil C.

; APPLICANT: Guegler, Karl J.

; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/884,072

; FILING DATE: Herewith

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0333 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 448 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: CORNNOT01

; CLONE: 45517

; US-08-884-072-1

Query Match

Best Local Similarity 51.1%; Score 1283; DB 2; Length 448;

Matches 225; Conservative 74; Mismatches 132; Indels 22; Gaps 5;

QY 8 LFGSLLWALLLLGASPDSEEPDSYTECTDGYEWDPDSDQHCRDNECLTPEACKG 67

Db 1 MGIKIRILVTTLALCLPSGNAQ-----AQTNFDLDRQSGCLDIDECRTIPEACRG 55  
QY 68 EMKCNHGYGGLCLPSRAAVINDLHG-----EGP-PPVPPAQNPN-----PCPP 111  
Db 56 DMWCNQNGGLCIPRTNVPYGRPSNYPSTPYSGPYPAAAPPLSAPNYPPTISRPLICRF 115  
QY 112 GYEPDQDSCVDVDECAQALHDCRPSQDCHNLPGSVOCTCPDGKYRKGIGPECVDIDECRYR 171  
Db 116 GYOMDESQCVDDVDECATDTHQCNPTQICINTEGGYTCSTDGYWLLEGQCLDIDECRYG 175  
QY 172 YQHQRCVNLPGSFRCOCEPGFQGLGNRRSCVDYNECDMGAPCEQRCFNSYGTFLCRCHOG 231  
Db 176 YCQQLCANVPGSYSTCNPFGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDFG 235  
QY 232 YELHROGFSDDIDECYSYLYCQYRCVNEPGRFSCHPGQYQALLATRLCQDIDECESG 290  
Db 236 YELEEDGVHCSDMDECSFSEFLCQHECVNQPGTYFCSPPGYILLDDNRSCQDINECEHR 295  
QY 291 AHQCEAQTQVNFHGGYRCVDTNRCVEPIQVSENRCCLCPASNPLCRQPSIVHRYMTI 350  
Db 296 NHTCNLQOQTCYNLQGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPFITLYRDMDV 355  
QY 351 TSERSVPADVFOIQATSYPGAYNAFOIRAGNSQGDYFIROINNVMFAMLVLRPVGTGR 410  
Db 356 VSGRSVPADIFOMQATRYPGAYYIFQIKSGNEGREGFYMQTGPISATLVMTRPKGP 415  
QY 411 YVLDLMTMNSLMSYRASSVLRITVFGAYTF 443  
Db 416 IQLDLMTMTVNTVINFRGSSVIRLRIYVSQYPF 448

RESULT 5  
US-09-212-168-1  
; Sequence 1, Application US/09212168  
; Patent No. 6303765  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/212,168  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/884,072  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0333 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 448 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: CORNNOT01  
; CLONE: 45517  
; US-08-884-072-1

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CORNNOT01
; CLONE: 45517

```

Query Match	51.18;	Score 1283;	DB 4;	Length 448;
Best Local Similarity	49.79;	Pred. No.6.3e-97;		
Matches 225;	Conservative 74;	Mismatches 132;	Indels 22;	Gaps 5;
QY	8	LPGSILLWALLLLLGSSADSPDSEEDPSYTECTGYEWDPSQCHRDVNECLTPEACKG	67	
DB	1	MPGKRLTWTITALCLSPGNAQ-----ACCTNGFDLDRSQSGCLDIDECRTPEACKG	55	
QY	68	EMKCNHNGYGYLCLPRSAVINDLHG-----EGP-PPVPVPAQHNP-----PCPP	111	
DB	56	DMVCNQNGGYLCIDPTNPVTRGYSNPYSTPGYPAAAPLSAPNTPTISRPLICRF	115	
QY	112	GYEPDDDCSDVDVDECAQALHDCRPSODCHNLPGSYOCTCDGYRKGPECDVIDECRYR	171	
DB	116	GYQMBESNCQVDVDECATDSHQCHPTQICINTEGYTCSTDGYWLEGGCLDIDECRVG	175	
QY	172	YCOHRCVNLPGSFRQCQBPFGQLGPNNRSCVDVNECDMGAPCQRCSYSGTFLCRCHOG	231	
DB	176	YQQLCANPGSYSCTCPNGPFTLNEDGRSCDVNECATENPCVQTCVNTGYFICRCDPG	235	
QY	232	YELHRDGFSCSDIDRCYSYSSYLQYRCVNEBGRSCHPCOGYOLL-ATRLCQDIDECESG	290	
DB	236	YELEDPGVHCSMDCSRSEFLQHECVNQBPVTYFCPPGYILLDDNRSCQDINECHRR	295	
QY	291	AHQCSAQCFNFGHYGCVDTNRCVEFYIQVSENRCPLCPASNPLCREPSSIVHYMTI	350	
DB	296	NHTCNLQQTCTYLNQGGFXCIDPICEEPLYLRISDNRCMCFAENPGCRDQPFITYRDMDV	355	
QY	351	TSERSVPADVFOIQATSYTPGAYNAFQIRAGNSOGDGYIRQINNVFAMLVLPVTPGRE	410	
DB	356	VSGRSVPADIFOMQATRYTPGAYYIFQIKSNEGREGYFMQTGPISATLVMTPIKGPRE	415	
QY	411	YVLDLEMTNMSLMYSRASSVLRTVFVGATTF	443	
DB	416	IOLDLEMTVNTVINFRGSSVIRLRYVSYOPF	448	

```

RESULT      6
US-08-884-072-5
; Sequence 5, Application US/08884072
; Patent No. 5872234
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,072
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

```

```

: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0333 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 387 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 458228
: US-08-884-072-5
:
: Query Match 45.6%; Score 1147; DB 2; Length 387;
: Best Local Similarity 55.7%; Pred. No. 6.3e-86;
: Matches 196; Conservative 61; Mismatches 87; Indels 8; Gaps 4;
:
Qy 99 PVPQAHP-NP-----CPPGYEPDDQSCVDVDECAALHDCRFSDQCHNLPGSYQCTCP 152
Db 37 PADPQRIPSNPSHRIQCAAGYEQSEHNVCQDIDECTAGTHNCRADQVCINLRGSFACQCP 96
Qy 153 DGYRKIGPECVDIDECRY-RYCOHRVCNVLGSPFCQCEPGFQLGPNNRSCVDVNECDMGA 211
Db 97 PGYQKRGECQVDIDECTIPPYCHORCVNTPGSFYQCQSPGQLAANNITCVDINECDASN 156
Qy 212 PCEQRCSFNSYGTFLCRCHQGHGYELHRDGFSCSDIDECYSYSLCYQRCVNEPGRFSCHCPQ 271
Db 157 QCAQCYNILGSGFCQCNGHGYELSDRLNCEIDECRTSSYLCYQCVNEPGRFSCHCPQ 216
Qy 272 GYQLLATFLCDDIDECRSAGHQCSEAQTCVNFHGGYRCVDTNRCVVEPIYQVSENRCLCPA 331
Db 217 GYQVVRSTCODINECBT-TNECREDEMCWNHGGFRICYPRNCPDPIYLTPENRVCVPV 275
Qy 332 SNPLCRPSPSIVHRYMTITSERVPADVFOIATSVYPGAYNAFQIRAGNSQDGFYIRQ 391
Db 276 SNAMCRELPQSVIVKYSIRSDRSPVSDIFOIQTATIVANTINFRKSGNENGFEYLQ 335
Qy 392 INNVMFLVLARPVTGPREYVLDLEWMTNMSYRASSVLRLRVFGAYTF 443
Db 336 TSPYSAMLVLVKSLSGPREHIVDLEMLTVSSIGTERTSSVLRLTIIVGPFSE 387

```

```

7
RESULT
US-08-833-963C-9
Sequence 9, Application US/08833963C
Patent No. 5916769
GENERAL INFORMATION:
APPLICANT: Olsen, et al.
TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
TITLE OF INVENTION: HCABA58X
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Ave
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,963C

```



```
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 435
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US96/05033
; FILING DATE: 10-APR-1996
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF258
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-963C-9

Query Match 45.6%; Score 1147; DB 2; Length 387;
Best Local Similarity 55.7%; Pred. No. 6.3e-86;
Matches 196; Conservative 61; Mismatches 87; Indels 8; Gaps 4;

QY 99 PVPAPQHP-NP-----CPPGYEPDDQSDVDDECAQALHDCRPSQDCHNLPGSYOCTCP 152
Db 37 PADPQRPSPHRIICAGYEQSEHNVCODIDECTAGTHNCRADQVCINLRGSFACQCP 96

QY 153 DGYRKIGPECVDIDECY-RYCQHRVCNLPSPFCQCEPGFQGLGNRSQVDVNECDMGA 211
Db 97 PGYQKGEQCVDIDECY-RYCQHRVCNLPSPFCQCEPGFQGLGNRSQVDVNECDMGA 156

QY 212 PCQRQFNSYGTFLCRHCHQGYELHRDGFSCSDIDEGSYSLQYRCVNPGRFSCHCPQ 271
Db 157 QCAQOCYNILGSPICQCNQGYELSLDNLNEDIDECRTSSYLQYQCVNPPGKFSQCMCPQ 216

QY 272 GYQLLATRLCQDIDEGSAGHOCSEATQCVNFGHGVRCVDTNRCVPEYIOVSENRLCLCPA 331
Db 217 GYQVRSRTQDINECT-TNECREDEMCWNYHGGFRCPYPRNCPQDPYILTPENRCVCPV 391

QY 392 INNVFAMVLARVPTGPREYVLDLEMTNLSMYRASSVLRITLVFGAYTF 443
Db 336 TSPVSAMVLVLSLGPRIHVLDLEMTNLSMYRASSVLRITLVFGAYTF 387

RESULT 8
US-08-980-514-3
; Sequence 3, Application US/08980514
; Patent No. 6004753
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN S1-5-ECMP-LIKE PROT
; TITLE OF INVENTION: EIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,514
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0436 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 458228
; US-08-980-514-3

Query Match 45.6%; Score 1147; DB 3; Length 387;
Best Local Similarity 55.7%; Pred. No. 6.3e-86;
Matches 196; Conservative 61; Mismatches 87; Indels 8; Gaps 4;

QY 99 PVPAPQHP-NP-----CPPGYEPDDQSDVDDECAQALHDCRPSQDCHNLPGSYOCTCP 152
Db 37 PADPQRPSPHRIICAGYEQSEHNVCODIDECTAGTHNCRADQVCINLRGSFACQCP 96

QY 153 DGYRKIGPECVDIDECY-RYCQHRVCNLPSPFCQCEPGFQGLGNRSQVDVNECDMGA 211
Db 97 PGYQKGEQCVDIDECY-RYCQHRVCNLPSPFCQCEPGFQGLGNRSQVDVNECDMGA 156

QY 212 PCQRQFNSYGTFLCRHCHQGYELHRDGFSCSDIDEGSYSLQYRCVNPGRFSCHCPQ 271
Db 157 QCAQOCYNILGSPICQCNQGYELSLDNLNEDIDECRTSSYLQYQCVNPPGKFSQCMCPQ 216

QY 272 GYQLLATRLCQDIDEGSAGHOCSEATQCVNFGHGVRCVDTNRCVPEYIOVSENRLCLCPA 331
Db 217 GYQVRSRTQDINECT-TNECREDEMCWNYHGGFRCPYPRNCPQDPYILTPENRCVCPV 275

QY 392 INNVFAMVLARVPTGPREYVLDLEMTNLSMYRASSVLRITLVFGAYTF 443
Db 336 TSPVSAMVLVLSLGPRIHVLDLEMTNLSMYRASSVLRITLVFGAYTF 387

RESULT 9
US-09-212-168-5
; Sequence 5, Application US/09212168
; Patent No. 6303765
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
```



OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,722B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US95/02251  
FILING DATE: 21-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/316,650  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/199,780  
FILING DATE: 18-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Fussey, Shelley P.M.  
REGISTRATION NUMBER: 39,458  
REFERENCE/DOCKET NUMBER: 4100.000500/FUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (713) 934-7000  
TELEFAX: (713) 934-7011  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1833 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-722B-2

Query Match 20.8%; Score 523.5; DB 3; Length 1833;  
Best Local Similarity 30.5%; Pred. No. 2.3e-34;  
Matches 128; Conservative 44; Mismatches 139; Indels 109; Gaps 19;  
QY 33 PDSY-TECTDGYEWDPSQHCVRDNECLTPEACKEMKINHYGVLCLPRSAVINL 91  
Db 876 PNGYRCVSPGYQLHPSQDYCTDDNECMRNP--CEGRGRCVNSVGSYSL----- 923  
QY 92 HGBGPPPPVPPAQAHPNCPGPGY---EPDQDSCVDVDEC----- 127  
Db 924 -----CYPGYTLVTLGDTQECQDIDECEQPGVCGSGRCNTEGSYHCE 966  
QY 128 -----AQAHLDCR-----PSQDCHNLPGSYQC--TCPDGYRKIGTPECVDIDEC- 168  
Db 967 CDGRGYIMVRKGGHQCQDINECRHPTGCPDGRVNSPGSYTCLACEGYVGSGSCVDVNECL 1026  
QY 169 RYRYCQH-RCVNLPGSFRQCEPGFOLGPNRNSVDVNECDMGAPCEQ-RCFNSYGTFLC 226  
Db 1027 TPGICHTHRCINNEGSRFCSEPGYEVTPDKKCRDVEDCASRASCTGLCLNTEGSFTC 1086  
QY 227 -RCHQGYELHRDGFSCSDIDECSYSLYLCQYRCVNEPGRFSC-HCPQGYQ--LLATRLCQ 282  
Db 1087 SACQSGYWNEDGTACEDLDECAFPVCTVNTVGSFCKDCQDQGYRPNPLGNR-CE 1145  
QY 283 DIDECSGAHQCEAQTQVNFHGGYRCV-----DTNRCVPEYIQVSENRL 328  
Db 1146 DVDECEGPQSSCRGGE-CKNTEGSYQCLCHQGFQVNGTMCEDVNECVGEEHCAPHECL 1204  
QY 329 CPASNPLCREQPSIVHYRYMTITRSVPA-DVFQIQATSVYPGAYNAFQIRAGNSQGD 387  
Db 1205 NSLGSFFCLCAPG-----FASAEGRTRCQDQVDECAATDPCPGH-----CVNTEGSF 1251

RESULT 12  
PCT-US95-02251-18  
Sequence 18, Application PC/TUS9502251  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE  
TITLE OF INVENTION: CELLS  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:

ADDRESSER: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: PatentIn Release #1.0, Version  
SOFTWARE: #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02251  
FILING DATE: CONCURRENTLY HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/316,650  
FILING DATE: 30-SEP-1994  
CLASSIFICATION:  
APPLICATION NUMBER: US 08/199,780  
FILING DATE: 18-FEB-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UMIC009P--  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1833 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-02251-18

Query Match 20.8%; Score 523.5; DB 5; Length 1833;  
Best Local Similarity 30.5%; Pred. No. 2.3e-34;  
Matches 128; Conservative 44; Mismatches 139; Indels 109; Gaps 19;  
QY 33 PDSY-TECTDGYEWDPSQHCVRDNECLTPEACKEMKINHYGVLCLPRSAVINL 91  
Db 876 PNGYRCVSPGYQLHPSQDYCTDDNECMRNP--CEGRGRCVNSVGSYSL----- 923  
QY 92 HGBGPPPPVPPAQAHPNCPGPGY---EPDQDSCVDVDEC----- 127  
Db 924 -----CYPGYTLVTLGDTQECQDIDECEQPGVCGSGRCNTEGSYHCE 966  
QY 128 -----AQAHLDCR-----PSQDCHNLPGSYQC--TCPDGYRKIGTPECVDIDEC- 168  
Db 967 CDGRGYIMVRKGGHQCQDINECRHPTGCPDGRVNSPGSYTCLACEGYVGSGSCVDVNECL 1026  
QY 169 RYRYCQH-RCVNLPGSFRQCEPGFOLGPNRNSVDVNECDMGAPCEQ-RCFNSYGTFLC 226  
Db 1027 TPGICHTHRCINNEGSRFCSEPGYEVTPDKKCRDVEDCASRASCTGLCLNTEGSFTC 1086  
QY 227 -RCHQGYELHRDGFSCSDIDECSYSLYLCQYRCVNEPGRFSC-HCPQGYQ--LLATRLCQ 282  
Db 1087 SACQSGYWNEDGTACEDLDECAFPVCTVNTVGSFCKDCQDQGYRPNPLGNR-CE 1145  
QY 283 DIDECSGAHQCEAQTQVNFHGGYRCV-----DTNRCVPEYIQVSENRL 328  
Db 1146 DVDECEGPQSSCRGGE-CKNTEGSYQCLCHQGFQVNGTMCEDVNECVGEEHCAPHECL 1204  
QY 329 CPASNPLCREQPSIVHYRYMTITRSVPA-DVFQIQATSVYPGAYNAFQIRAGNSQGD 387  
Db 1205 NSLGSFFCLCAPG-----FASAEGRTRCQDQVDECAATDPCPGH-----CVNTEGSF 1251







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2002, 09:38:45 ; Search time 20.87 Seconds  
(without alignments)  
2039.654 Million cell updates/sec

Title: US-09-829-936A-22  
Perfect score: 2513  
Sequence: 1 MLCASCLPGSLLWALLLL.....MSYRASSVLRLTVFGAYTF 443  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1348	53.6	493	2 JC5621	epidermal growth f
2	1147	45.6	387	2 I38449	extracellular prot
3	805	32.0	1221	2 A49457	fibulin-2 precursor
4	796.5	31.7	1184	2 A55184	fibulin-2 precursor
5	747	29.7	685	2 S78040	fibulin, splice fo
6	711	28.3	683	2 C36346	fibulin 1 precursor
7	698.5	27.8	705	2 S34968	fibulin, splice fo
8	596	23.7	601	2 B36346	fibulin 1 precursor
9	555.5	22.1	689	2 T42760	fibulin, splice fo
10	555	22.1	589	2 T43210	fibulin-1D precurs
11	549.5	21.9	712	2 T42990	fibulin 1, splice
12	529.5	21.1	798	2 T22793	hypothetical prote
13	523	20.8	2918	2 A54105	fibillin-2 precur
14	518.5	20.6	2907	2 A57278	fibillin-2 precur
15	511.5	20.4	2871	2 A55567	fibillin I - bovi
16	508.5	20.2	3002	2 A47221	fibillin 1 precur
17	507.5	20.2	2871	2 A55624	fibillin-1 precur
18	502.5	20.0	1820	2 A55494	latent transformin
19	480	19.1	1394	2 A35626	transforming growt
20	476.5	19.0	1712	2 A38261	masking protein pr
21	452	18.0	1251	2 A57293	latent transformin
22	451	17.9	1620	2 T27283	hypothetical prote
23	441.5	17.6	1574	2 T13954	MEGF6 protein - ra
24	440	17.5	3507	2 T34513	hypothetical prote
25	438	17.4	741	2 T46488	hypothetical prote
26	403.5	16.1	2555	2 A40043	notch protein homo
27	399	15.9	2531	2 S18188	notch protein homo
28	388	15.4	886	2 A57172	probable hormone r
29	380	15.1	2524	2 A35844	Xotch protein - Af

RESULT 1

JC5621  
epidermal growth factor-like protein, T16 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 09-Oct-1997 #sequence\_revision 07-Nov-1997 #text\_change 05-Nov-1999  
C:Accession: JC5621  
R:Ozaki, T.; Kondo, K.; Nakamura, Y.; Ichimiya, S.; Nakagawara, A.; Sakiyama, S.  
Biochem. Biophys. Res. Commun. 237, 245-250, 1997  
A:Title: Interaction of DA41, a DAN-binding protein, with the epidermal growth factor  
A:Reference number: JC5621; MUID:97415782  
A:Accession: JC5621  
A:Molecule type: mRNA  
A:Residues: 1-493 <OZA>  
A:Cross-references: DBBJ:D89730; NID:g2429082; PIDN:BAA22265.1; PID:d1023127; PID:g24  
C:Comment: This protein plays a role in the regulation of cell growth by interacting  
C:Keywords: glycoprotein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:28-70,158-199,200-237,238-277,278-318,319-359/Region: epidermal growth factor-like  
F:249/Binding site: carbohydrate (Asn) (covalent) #status predicted

ALIGNMENTS

Query Match	53.6%	Score	1348	DB	2	Length	493
Best Local Similarity	48.0%	Pred. No.	2.8e-82				
Matches	235	Conservative	76	Mismatches	113	Indels	66
Gaps	4						
QY	17	LLLLLIGSASQDSEPPDSYECTDGYEWDPSQHQRDVNECLTPEACKGEMKCNHYG	76				
Db	7	LTMLTALVKSQVTEITTYTCTDGYEWDVPRQCKDIDCEDIVPDACKGGMKCVNHVG	66				
QY	77	GYLCPLRSAVINLHGEPPPPVPPAQ-----	104				
Db	67	GYLCPLPKTAQII--VNNEQPQQTAAEAASGAATGTAARSMAATSGVTPGGFIASATA	124				
QY	105	-----HNPCCPGYEPDODSDVDDECAQLHDC	134				
Db	125	VAGPEVQTGRNNFVTRNPADPQIRPSNHRITQCAAGYEQSEHNVQCQIDICTSGTHNC	184				
QY	135	RFSDQCHNLPGSYQCTCPDGYKRGKIGPECVDIDECRY-RYQHRVCNLPDGSFRQCQBPFGQ	193				
Db	185	RLDQVCINLNGSFTCHCLGYQKRGQCVDIDECSPVPYCHQCCVNTPGSFYCCQNPFGQ	244				
QY	194	LGPNNRSCVDVNECDMGAPCEQRCFNSTGYTFLCRCHQGYELHRDGFSCSDIDECSSSYL	253				
Db	245	LAANNVCVDINECDASNOCAOCCNYLGSFCQCCNOGYELSSDRLNCEDIDECRTSSYL	304				
QY	254	COYRCVNEGRSCHCPQGYQLLATRLCQDIDECESGAHQCEAQTCTVNFHGRCYRVDIN	313				
Db	305	COYQCVNEPGRFTSCMCPQGYQVVRSTCQDINECET-TNECREDECMWYHGGFCYPON	363				
QY	314	RCVEPYIQVSENRCILCPASNPLCREQPSIVHRYMTITTSERSVPADVFOIQTATSVVPGAY	373				
Db	364	PCQDPYVLTSENRCVCPVSNWTCRDPQPSIVYKMYMIRSDRSPVDFIQIQTATYANTI	423				

Notch-1 protein -  
transmembrane prot  
cell-fate determin  
notch protein - fr  
notch B protein -  
notch 3 protein -  
notch homolog - se  
notch3 protein - h  
notch4 - mouse  
hypothetical prote  
growth potentiati  
crumbs protein - f  
growth arrest-spec  
growth arrest-spec  
Notch homolog Motc  
Nel-homolog protei









QY 98 PPVPPAHPNCPGYPEDDQDS-CVDVDECAQALHDCRPSQDCHNLPGSYOCTCPDGY- 155  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 335 ----QKNVFN-CGRGVHLEEGTRCVDVDECAAPPABPCGKHRCVNSPGSFRCECKTGY 389  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 156 -RKIGPECVDIDEC-RY--RYCOHRCVNLPGSFRCOCEPGFOLGPNRSCVDVNECDMA 211  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 390 FDGISMCDVNECQYRGLCHCKENTLGSYLSCSVGFLSVSDGRSCEDINECS-SS 448  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 212 PCBQCFNSYGTFLCRCHGYELHR-DGFSQDIDECYSY--SSYLQYRCVNEPGRFSCH 268  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 449 PCSQECANVGSYQCYCRGQYQLSDVGYTCEDIDECALPTGGHICSYRINIPGSFQCS 508  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 269 CP-OGVOLLAT-RLCQDIDECSSGAHQCEAQTGVNPHGGRYCDVTRNCEVPY 319  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 509 CPSSGYRLAPNGRNCQDIDECVTGHNCSINETCFNIOGAFCRL-AFCECPENY 560  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |

RESULT 9  
T42760  
fibulin, splice form D precursor - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C:Accession: T42760  
R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.  
submitted to the EMBL Data Library, February 1998  
A:Description: Isolation of chicken and nematode fibulin-1 homologs and characterization  
A:Reference number: Z22267  
A:Accession: T42760  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-689 <BAR>  
A:Cross-references: EMBL:AF051401; PIDN: AAC28321.1  
C:Genetics:  
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 22.1%; Score 555.5; DB 2: Length 689;  
Best Local Similarity 26.4%; Pred. No. 1.3e-29;  
Matches 144; Conservative 67; Mismatches 158; Indels 177; Gaps 22;

QY 39 CTGGEWDPSQHCQDVNECLTIPEACKGEMKCIHYGYLCLPRSAVINLHGEPPPP 98  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 180 CRSGFOLAPDGMACVDIDECATLMDCLSQRCNLNTPGSFKCI-RTLS ----- 226  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 99 PVPPAHPNCPGYPEDDQ-DSQVDVDECAQALHDCRPSQDCHNLPGSYOCTCPDGY- 149  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 227 -----CGTGYAMDSETERCRDVECLNGLSHDCGPLYQCRNTQGSYRCDACKKGDG 276  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 150 -----TCPDGYRKIGPECVDIDEC-----RYRYC 173  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 277 ELONPMTGECTSITCPNGYYPKNGMNDIDECVTGHNCGAGEECVNTPGSFRQQRGNLC 336  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 174 QH-----RCVNLPL----- 181  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 337 AHGYENYGATGFCEDYNECQGVCGSMECINLPGTYKCKGPGYEFNDACKKRCEDVDECI 396  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 182 -----GSFRCOCEPGFOLGPNRSCVDVNECDMG-APCEQRCFNSYGT 223  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 397 KFAGHVCDLSAECINTIGSFCKKPGFOLASDGRRCEDVNECTTGIAACEQKCVNIPGS 456  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 224 FLCRCHQGYELHRDGFSCSDIDECYSY-----SSYLQYRCVNEPGRFSCHCPQYOLLAT- 278  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 457 YQCICDRGAFALPDGPKCEDIDECSTWAGSNDLCMGCCINTKGSYLCCQCPGYKTOPDG 516  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 279 RLCQDIDECESGAHQCEA-OTCVNPHGGRYCDVTRNCEVPYIOVSENK-----CL- 328  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 517 RTCVDVDECAMG--ECAGSKVCVNTLGSFKCHSID-CPTNYIHDSLKNQIADGSCIK 573  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 329 -CPASNPLC-REQPSSIVHRYMTITSERV--PADVFOIQATSVYPGA--YNAFQIRAGN 382  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 574 VCSTEDTECLGNHTRVLYQFRAVPSLKTIIISPIEVSRIVTHMGVPFSDYNLDYV---- 629  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |

QY 383 SQGDFVIRQINNVFAMLVLARPVGTGPREVYLDLMTWMTNLSMSYRASSVLR-----LTVF 437  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 630 GORHFRIVQERNI-GIVQLVKPISGP-----TVETIKVNIHTKSRTGVILAFNEALIEIS 683  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 438 VGAYTF 443  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 684 VSKYPF 689  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |

RESULT 10  
T43210  
fibulin-1D precursor - Caenorhabditis elegans (fragment)  
C:Species: Caenorhabditis elegans  
C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T43210  
R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.  
submitted to the EMBL Data Library, June 1998  
A:Description: Identification of chicken and C. elegans fibulin-1 homologs and charac  
A:Reference number: Z22337  
A:Accession: T43210  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-589 <BAR>  
A:Cross-references: EMBL:AF070477; PIDN: AAC24035.1  
C:Genetics:  
A:Note: Intron positions not resolved (incomplete sequence)  
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 22.1%; Score 555; DB 2: Length 589;  
Best Local Similarity 29.1%; Pred. No. 1.2e-29;  
Matches 144; Conservative 68; Mismatches 162; Indels 120; Gaps 25;

QY 39 CTGGEWDPSQHCQDVNECLTIPEACKGEMKCIHYGYLCLPRSAVINLHGEPPPP 98  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 127 CGTGYAMDSETERCRDVECLNGLSHDCGPLYQCRNTQGSYRCDACKK-----GDG--- 176  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 99 PVPPAHPNCPGYPEDDQDSQVDVDECAQALHDCRPSQDCHNLPGSYOCTCPDGY- 149  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 177 -----ELONPMTGECTSITCPNGYYPKNGMNDIDECVTG-HNCGAGEECVNTPGSFR 229  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 150 -----TCPDGYR-----KIGP-----EC 162  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 230 QOKGNLCAHGYENYGATGFCEDYNECQGVCGSMECINLPGTYKCKGPGYEFNDACKKRC 289  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 163 VDIDECYRYCQH-----RCVNLPGSFRCOCEPGFOLGPNRSCVDVNECDMG-APCEQ 215  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 290 EDVDEC-IFAGHVCDLSAECINTIGSFCKKPGFOLASDGRRCEDVNECTTGIAACEQ 348  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 216 RCFNSYGTFLCRCHQGYELHRDGFSCSDIDECYSY-----SSYLQYRCVNEPGRFSCHCPQ 271  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 349 KCVNIPGSYQCICDRGAFALPDGPKCEDIDECSTWAGSNDLCMGCCINTKGSYLCCQCP 408  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 272 GYOLLAT-RLCQDIDECESGAHQCEA-OTCVNPHGGRYCDVTRNCEVPYIOVSENK----- 326  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 409 GYKIQPDGRTCDVVDDECAMG--ECAGSKVCVNTLGSFKCHSID-CPTNYIHDSLKNQI 465  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 327 -----CL--CPASNPLC-REQPSSIVHRYMTITSERV--PADVFOIQATSVYPGA--YN 374  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 466 ADGYSCKIKVCSTEDTECLGNHTRVLYQFRAVPSLKTIIISPIEVSRIVTHMGVPFSDYN 525  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 375 AFQIRAGNSQGDYFIQINNVFAMLVLARPVGTGPREVYLDLMTWMTNLSMSYRASSVLR- 433  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 526 LDYV-----GORHFRIVQERNI-GIVQLVKPISGP-----TVETIKVNIHTKSRTGVILAF 575  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
QY 434 ----LTVFVGAYTF 443  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |  
Db 576 NEALIEISVSKYPF 589  
| | | | | : : : : | | | | | : : : : | | | | | : : : : | |

RESULT 11  
T42990

fibulin 1, splice form C precursor - *Caenorhabditis elegans*  
 C/Species: *Caenorhabditis elegans*  
 C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
 C/Accession: T42990  
 R/Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.  
 submitted to the EMBL Data Library, February 1998  
 A/Description: Isolation of chicken and nematode fibulin-1 homologs and characterization  
 A/Reference number: Z22267  
 A/Accession: T42990  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-712 <BAR>  
 A/Cross-references: EMBL:AF051402; PIDN:AAC28322.1  
 C/Genetics:  
 A/Gene: FBLN1  
 C/Superfamily: unassigned EGF-related proteins; EGF homology  
 C/Keywords: alternative splicing; basement membrane; extracellular matrix

Query Match	21.9%;	Score 549.5;	DB 2;	Length 712;
Best Local Similarity	26.5%;	Pred. No. 3.4e-29;		
Matches 141;	Conservative	61;	Mismatches 137;	Indels 193;
				Gaps 20;

QY	39	CTDGYEWDPDSQHCRDQVNECLTTPERACKGEMKCNINHYGYLCPLRSGAAVINDLHGSGPPP	98
Db	180	CRSGFOLAPDGMACVDDIECATLMDCLSDLESQRLNTPGSKFCT-RTLS	226
QY	99	PVPPAQHPNCPGYPEDDQ--DSCVDVDECAQALHDCRPSQDCNHLPGSVQC-----	149
Db	227	-----CGTGYAMDSETERCRDVEDCNLGHSDCGPLYQCRNTQGSYRCDAKKCGD	276
QY	150	-----TCPDGYRKIGPECVDDIEC-----RYRYC	173
Db	277	ELQNPMTGECTSITCPNGYYPFKMGMCNDIDECVTGHNCGAGECVNTPGSRFCQQKGNLC	336
QY	174	QH-----RCVNLPL-----	181
Db	337	AHGYEVNGATGCFEDVNECOQVCGGSMWCINLPGTYKCKGPGYFNDAKKRCEDYDECI	396
QY	182	-----GSRFCOCEPGFQLGPNNRSCVDVNECDMG-APCEQRCFNSYGT	223
Db	397	KFAGHVCDLSAECINTIGSFCECKPGFQLASDGRRCEDNECTTIGTAAECQKCVNIPGS	456
QY	224	FLCRHOGYELHRDGFSCSDIDECSY-----SSVLCQYRCVNEPGRFSCHCPGYQLLAT-	278
Db	457	YQCICDRGFALGPDGTYCKCEDIDECSTWAGSGNDLGMGGCINTHKGSLCOCPPGYKLPDGG	516
QY	279	RLCQIDIDECBSGAHQCSA-OTCNFPHGGYCYVDTRNCRPEPYIQVSENRCLC---PASN	334
Db	517	RTCVDVDDECAAG--ECAGSDKVCVNTLGSFKCHSID-CPTNYIHDLSLNKRNCRNQPSACG	573
QY	335	L---CREOPSSIVHRVYNTITSERSVP-----ADV---FQIQATSVYP	370
Db	574	LPECSKVPLFYVQFTSLA--RAVPLSSHRPAITLFKVSAPNHADTEFNVFQLQKTTIV	631
QY	371	GAYNA-----FQIRAGNSQGGDFYIRQINNVEFAMLVLARPTVGTGREYVLDL	415
Db	632	GAPNVLPAIRANFLLQKGERNS-----AVVTLRDSLDGPTQVKLOL	673

```

RESULT 12
T22793
hypothetical protein F56H11.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 13-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T22793; T24489
R:Wilkinson, J.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19616
A:Accession: T22793
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-798 <WIL>

```

A; Cross-references: EMBL:Z68749; PIDN:CAA92962.1; GSPDB:GN00022; CESP:F56H11.1  
A; Experimental source: clone F56H11  
R; Lloyd, C.  
submitted to the EMBL Data Library, December 1995  
A; Reference number: Z19897  
A; Accession: T24489  
A; Status: preliminary; translated from GB/EMBL/DBDJ  
A; Molecule type: DNA  
A; Residues: 1-798 <W12>  
A; Cross-references: EMBL:Z68219; PIDN:CAA92483.1; GSPDB:GN00022; CESP:F56H11.1  
A; Experimental source: clone T05A1  
C; Genetics:  
A; Gene: CESP:F56H11.1  
A; Map position: 4  
A; Introns: 14/1; 92/3; 144/1; 195/1; 281/1; 325/2; 371/2; 390/1; 437/1; 471/3;  
C; Superfamily: unassigned EGF-related proteins; EGF homology

Query Match	21.1%	Score 529.5	DB 2:	Length 798	
Best Local Similarity	24.5%	Pred. No. 7.9e-28			
Matches	146	Conservative 61	Mismatches 154	Indels 235	Gaps
QY	39	CTDGYEWDPSQSHCRDNECLTIPEACKGEMKCNHGYGLCLPRSAAVINDLHGSGPPP	98		
Db	180	CRSGFDLAPDMACVDNRECLTRQSPCTQSEDVCNTIGTICORRISRLVPHHRRNRITG	239		
QY	99	PVPPAQHPN-----CPPGYEPDQDSCVDVDCACALHDCRPSQDC	140		
Db	240	NAPRRMDDPYSRAGEYREASQANTEFGCPMGW-LFOHGHCVDIDECATLMDDCLSESQR	298		
QY	141	HNLPGSYQC-----TCPDGYRK-----	157		
Db	299	LNTPGSEKCIRTLSCGTGYAMDSETERNNCFLLIILNNTFNCKYFFVEDVDECNLGHSDGC	358		
QY	158	-----IGPECV-----	163		
Db	359	PLYQCRNTQSYRCDAKKGDGDELQNPMTGEYIDECVTGHNCGAGECVCNTPGSPFCQOK	418		
QY	164	-----DIDECRYRYC-OHRCVNLP-----	181		
Db	419	GNLCAHGYEVNGATGFCEDYNECQQGVCGSGMECINLPGYTKCKGPGYBFNDAKKRCEDV	478		
QY	182	-----GSFRCCQEPGFOLGPNNRSCVDVNECDMG-APCEORCFEN	219		
Db	479	DECIKFAHGVCDLSABECINTIGSFCKCKPGFOLASDGRRCEDVNECTTGIAACEOKCVN	538		
QY	220	SYGTFLLCRHQGYELHRDGFSCSIDBESY-----SSYLQYRCVNPFGFRFSCHCPGYQL	275		
Db	539	IPGSYQICIDRGFALGPDGKTKEDIDECISIWAGSGNDLGMGGCINTKGSYLQCQPPGYKI	598		
QY	276	LAT-RLCQDIDECESGAHQCEA-QTCVNPFGHYGRVDVNRNRCVEPIQVSENRCIC---	330		
Db	599	QPDGRTCDVDYDECAMG--ECAGSDKVCVNTLGSFKCHSID-CPTNTIHPSLAKNRNCRQP	655		
QY	331	ASNPL-----CREOPSSIVHRYMTTTSERSVP-----ADV---FOIQAT	366		
Db	656	SACGLPECSKVPLFTYQFISIA--RAVDISSHRPAITLFKVSAFNHADTEVNFELQLK	713		
QY	367	SVYPGAYNA-----FQIRAGNSQGDFFYIRQINNVFAMLVLARPYTGPREYVLDL	415		
Db	714	TTIIGAPNVLPAIRANFLLOKGEKRN-----AVVTLRDSLDGPTVKLOL	759		

```

RESULT 13
A54105
fibrillin-2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 17-Nov-2000
C:Accession: A54105; S17063; S31101
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio,
J. Cell Biol. 124, 855-863, 1994
A:Title: Structure and expression of fibrillin-2, a novel microfibrillar co
A:Reference number: A54105; MUIP:94165150

```



```
Db 2107 FRICPYGSGIIVGPDSDSAYDMDECKEPDVCKHGQCINTDGSYRCECPFGYILOGN--EC 2164
QY 202 VDVNECDMGAPC-EQRCFNYSYGTFLCRCHQGYELHRDG--FSCSDIDECSYSSYLQOYRC 258
Db 2165 VDTDECSVGNPCNGTCKNVIGGFECTCEBGF--PGPMITCEDINECAQNPFLCAFC 2221
QY 259 VNEPGRFSCPCOGYQLLA-TRLCQDIDECESGAHOCSEAQ-TCVNFHG-----GYR 308
Db 2222 VNTYGYECKCPAGYVLRDRRMCKDEDECEGKHDCAEKQMECKNLIGTYLCICGPGYQ 2281
QY 309 -----CVDTRNC-VEPYIOVSENRCICPASNPLC 336
Db 2282 RRPDGECDVNECQTKPGI-CENGRCLNTRGSYTC 2316
```

Search completed: September 25, 2002, 09:43:05  
Job time: 260 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2002, 09:42:40 ; Search time 13.41 Seconds  
(without alignments)  
1279.102 Million cell updates/sec

Title: US-09-829-936A-22  
Perfect score: 2513  
Sequence: 1 MLPSCSLPLGSLLLWALLL.....MSYRASSVRLTVFVGATP 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2505	99.7	443	1 FBL4_HUMAN	O95967 homo sapien
2	2445	97.3	443	1 FBL4_CRIGR	O55058 cricetus
3	2392	95.2	443	1 FBL4_MOUSE	O9wv19 mus musculus
4	1379	54.9	493	1 FBL3_HUMAN	Q12805 homo sapien
5	1348	53.6	448	1 FBL3_RAT	O35568 rattus norv
6	1283	51.1	448	1 FBL5_HUMAN	O9ubx5 homo sapien
7	1269	50.5	448	1 FBL5_MOUSE	O9wv19 mus musculus
8	1261	50.2	448	1 FBL5_RAT	O9wv18 rattus norv
9	805	32.0	1221	1 FBL2_MOUSE	P37889 mus musculus
10	796.5	31.7	1184	1 FBL2_HUMAN	P98095 homo sapien
11	747	29.7	684	1 FBL1_CHICK	O73775 gallus gall
12	698.5	27.8	705	1 FBL1_MOUSE	Q08879 mus musculus
13	664.5	26.4	703	1 FBL1_HUMAN	P23142 homo sapien
14	549.5	21.9	712	1 FBL1_CAEEL	O77469 caenorhabdi
15	524.5	20.9	2911	1 FBL2_HUMAN	P35556 homo sapien
16	518.5	20.6	2907	1 FBL2_MOUSE	O61555 mus musculus
17	511.5	20.4	2871	1 FBL1_BOVIN	P98133 bos taurus
18	509.5	20.3	2871	1 FBL1_PIG	O9tv36 sus scrofa
19	508.5	20.2	2871	1 FBL1_HUMAN	P35555 homo sapien
20	507.5	20.2	2871	1 FBL1_MOUSE	O61554 mus musculus
21	480	19.1	1394	1 TGFB_HUMAN	P22064 homo sapien
22	476.5	19.0	1712	1 TGFB_RAT	Q00918 rattus norv
23	436	17.3	956	1 MTN2_HUMAN	O00339 homo sapien
24	413.5	16.5	956	1 MTN2_MOUSE	O08746 mus musculus
25	399	15.9	2531	1 NTCL1_RAT	Q07008 rattus norv
26	393	15.6	2444	1 NTCL1_HUMAN	P46531 homo sapien
27	388	15.4	886	1 FBL1_HUMAN	Q14246 homo sapien
28	388	15.4	2524	1 NOTC1_XENLA	P21783 xenopus lae
29	376	15.0	2531	1 NTCL1_MOUSE	Q01705 mus musculus
30	373	14.8	2437	1 NOTC1_BRARE	P46530 brachydanio
31	360.5	14.3	816	1 NEL_CHICK	O90827 gallus gall
32	360.5	14.3	931	1 FBL1_MOUSE	O61549 mus musculus
33	359.5	14.3	816	1 NEL2_MOUSE	Q61220 mus musculus

ALIGNMENTS

RESULT 1

ID	FBL4_HUMAN	STANDARD;	PRT;	443 AA.
AC	O95967; O75967;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	EGF-containing fibulin-like extracellular matrix protein 2 precursor (Fibulin-4) (FBL-4) (UHL protein).			
DE	(Fibulin-4) (FBL-4) (UHL protein).			
GN	EFEMP2 OR FBLN4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Melanoma;			
RX	MEDLINE=20068041; PubMed=10601734;			
RA	Giltay R., Timpl R., Kostka G.;			
RT	"Sequence, recombinant expression and tissue localization of two novel extracellular matrix proteins, fibulin-3 and fibulin-4.";			
RL	Matrix Biol. 18:469-480(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Zemel R., Shaul Y.;			
RN	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20435063; PubMed=10982184;			
RA	Katsanis N., Venable S., Smith J.R., Lupski J.R.;			
RT	"Isolation of a paralog of the Doyme honeycomb retinal dystrophy gene from the multiple retinopathy critical region on 11q13.";			
RL	Hum. Genet. 105:66-72(2000).			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; AJ132819; CAA10791.2; -			
DR	EMBL; AF093119; AAC62108.1; -			
DR	EMBL; AF109121; AAF65188.1; -			
DR	HSSP; P35555; 1EMN.			
DR	MIM; 604633; -			
DR	InterPro; IPR000152; Asx_hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	InterPro; IPR001491; Thrbomodulin.			
DR	Pfam; PF00008; EGF_4.			
DR	PRINTS; PR00907; THRBOMODULN.			
DR	SMART; SM00179; EGF_CA; 4.			

SMART; SM00001; EGF\_like; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 4.  
DR PROSITE; PS01187; EGF\_CA; 6.  
KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.  
FT SIGNAL 1 25  
FT CHAIN 26 443  
FT  
FT  
FT DOMAIN 36 81  
FT DOMAIN 123 163  
FT DOMAIN 164 202  
FT DOMAIN 203 242  
FT DOMAIN 243 282  
FT DOMAIN 283 328  
FT DOMAIN 329 368  
FT DISULFID 127 140  
FT DISULFID 134 149  
FT DISULFID 151 162  
FT DISULFID 168 177  
FT DISULFID 173 186  
FT DISULFID 188 201  
FT DISULFID 207 217  
FT DISULFID 213 226  
FT DISULFID 228 241  
FT DISULFID 247 258  
FT DISULFID 254 267  
FT DISULFID 269 281  
FT DISULFID 287 300  
FT DISULFID 294 309  
FT DISULFID 315 327  
FT CARBOHYD 198 198  
FT CARBOHYD 394 394  
FT CONFLICT 5 5  
FT CONFLICT 44 51  
FT CONFLICT 103 111  
FT CONFLICT 294 294  
FT CONFLICT 354 356  
FT CONFLICT 355 355  
SQ SEQUENCE 443 AA; 49391 MW; 9E9AC2393780D3B8 CRC64;

Query Match 99.7%; Score 2505; DB 1; Length 443;  
Best Local Similarity 99.8%; Pred. No. 3.7e-179;  
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLPASCLPGSLLLWALLLLLSASQDSEPDSTECTDGYEWDPSQHC RDVNECLT 60  
DB 1 MLPASCLPGSLLLWALLLLLSASQDSEPDSTECTDGYEWDPSQHC RDVNECLT 60  
QY 61 IPEACKGEMKCNHGYGLCPRAAVIDLHGEGPPPPVPPAHPNCPGPEPDQDS 120  
DB 61 IPEACKGEMKCNHGYGLCPRAAVIDLHGEGPPPPVPPAHPNCPGPEPDQDS 120  
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCOHRCVNL 180  
DB 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCOHRCVNL 180  
QY 181 PGSPRCQCEPGFQGLPNRRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240  
DB 181 PGSPRCQCEPGFQGLPNRRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240  
QY 241 CSDIDECSSYLCOYRCVNEPGRFSCPCPGYQLLATRLCQDIDECESAHQCSAQC 300  
DB 241 CSDIDECSSYLCOYRCVNEPGRFSCPCPGYQLLATRLCQDIDECESAHQCSAQC 300  
QY 301 VNFHGGYRCVDTNRCVPEYIQVSENRCCLPASNPCLREQSPSSIVHRYMTTTSERSPADV 360  
DB 301 VNFHGGYRCVDTNRCVPEYIQVSENRCCLPASNPCLREQSPSSIVHRYMTTTSERSPADV 360  
QY 361 FQIQATSVYPCAYNAFOIRAGNSQGDYFIRQINNVMFAMVLARVTPGPREYVLDLEMTM 420  
DB 361 FQIQATSVYPCAYNAFOIRAGNSQGDYFIRQINNVMFAMVLARVTPGPREYVLDLEMTM 420

QY 421 NSLMSYRASSVLRLLTVFGAYTF 443  
DB 421 NSLMSYRASSVLRLLTVFGAYTF 443  
RESULT 2  
FBL4\_CRIGR  
ID FBL4\_CRIGR STANDARD; PRT; 443 AA.  
AC O5058;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE EGF-containing fibulin-like extracellular matrix protein 2 precursor  
(Fibulin-4) (FBLN-4) (H411 protein).  
GN EFEMP2 OR FBLN4.  
OS Cricetulus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetulus.  
OX NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Heine H., Delude R.L., Monks B., Golenbock D.T.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@sib-sib.ch).  
CC -----  
DR EMBL; AF046870; AAC03101.1; -  
DR HSP; P00736; IAPQ.  
DR InterPro; IPR000152; ASX\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_CA.  
DR Pfam; PF00008; EGF; 4.  
DR SMART; SM00179; EGF\_CA; 4.  
DR SMART; SM00001; EGF\_like; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 4.  
DR PROSITE; PS01187; EGF\_CA; 6.  
KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.  
FT SIGNAL 1 25  
FT CHAIN 26 443  
FT  
FT DOMAIN 36 81  
FT DOMAIN 123 163  
FT DOMAIN 164 202  
FT DOMAIN 203 242  
FT DOMAIN 243 282  
FT DOMAIN 283 328  
FT DISULFID 127 140  
FT DISULFID 134 149  
FT DISULFID 151 162  
FT DISULFID 168 177  
FT DISULFID 173 186  
FT DISULFID 188 201  
FT DISULFID 207 217  
FT DISULFID 213 226  
FT DISULFID 228 241  
FT DISULFID 247 258  
FT DISULFID 254 267  
FT DISULFID 269 281  
FT DISULFID 287 300  
FT DISULFID 294 309  
FT DISULFID 315 327







Qy	194	LGPNNRSCVDVNECDMGAPCEQRCFNSYGTFFLCRCHQGYELHRDGFSCSDIDECYSYSSL	253
Db	245	LAANNYTCVDINECDSNACQCYNLGSGFIQCQNGGYELSSDRULIDECRTSYSL	304
Qy	254	COYRCVNEPGRFSCHPQGGYQLLATRLCODIDECSEGAHQCEAPQCVNFHGGRVCDIN	313
Db	305	COYQCVNEPGRFSCHPQGGYQVYRSTCODINECET-TNECREDEMCWNIHGGRFCYPRN	363
Qy	314	RCVEPYIQVSENCLCPASNPCLREQPSIVHRYMTITSESRVPADVFOIQATSYVPGAY	373
Db	364	PCQDPYILTLPENRCVCPVSNAMCRELPSQIVYKYSIRSDRSPVSDIFQIQATTIVANT	423
Qy	374	NAFOIRAGNSOGDFYIROINNVFAMLVLLARPVTGPREYVLDEMTNMSLMYRASSVL	433
Db	424	NTFRKSGNNGEYFLKQTSFVSAMLVLVKSLGGPREHIVDLEMLTVSSIGTFRTSVLR	483
Qy	434	LTVFVGAYTF	443
Db	484	LTIVGPFSE	493

RESULT	5
ID	FBL3_RAT
FB_L3	RAT
STANDARD;	PRT; 493 AA.
AC	Q35568;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	EGF-containing fibulin-like extracellular matrix protein 1 precursor
DE	(Fibulin-3) (FBL3-3) (Tl6 protein).
GN	EFEMP1 OR FBLN3.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX	NCBI_TaxID=10116;
ON	[1]
RN	SEQUENCE FROM N.A.
RP	TISSUE=Lung;
RC	MEDLINE=97415782; PubMed=9268694;
RX	Ozaki T., Kondo K., Nakamura Y., Ichimiya S., Nakagawara A.,
RA	Sakiyama S.;
RT	"Interaction of DA41, a DAN-binding protein, with the epidermal growth
RT	factor-like protein, s(1-5).";
RL	Biochem. Biophys. Res. Commun. 237:245-250(1997).
CC	-I- SUBCELLULAR LOCATION: Secreted.
CC	-I- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collabor-
CC	between the Swiss Institute of Bioinformatics and the EMBL outpost
CC	the European Bioinformatics Institute. There are no restrictions on
CC	use by non-profit institutions as long as its content is in no
CC	modified and this statement is not removed. Usage by and for commo-
CC	entities requires a license agreement (See http://www.isb-sib.ch/anno-
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: D89730; BAA22265.1; -
DR	HSP: P35555; 1EMN
DR	InterPro: IPR000152; Asx_hydroxyl.
DR	InterPro: IPR000561; EGF-like.
DR	InterPro: IPR001881; EGF_Ca.
DR	Pfam: PF00008; EGF_3.
DR	SMART: SM00179; EGF_CA; 4.
DR	SMART: SM00001; EGF_like; 2.
DR	PROSITE: PS00010; ASX_HYDROXYL; 4.
DR	PROSITE: PS00022; EGF_1; FALSE_NEG.
DR	PROSITE: PS01186; EGF_2; 4.
DR	PROSITE: PS01187; EGF_CA; 6.
KW	Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT	SIGNAL 1 POTENTIAL.
FT	CHAIN 18 493 EGF-CONTAINING FIBULIN-LIKE EXTRACELLULAR
FT	MATRIX PROTEIN 1.
FT	FT EGF-LIKE 1, DIVERGENT.
FT	DOMAIN 26 71

FT	DOMAIN	173	213	EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	214	253	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	254	293	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	294	333	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	334	373	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT	DISULFID	177	199	BY SIMILARITY.
FT	DISULFID	184	199	BY SIMILARITY.
FT	DISULFID	201	212	BY SIMILARITY.
FT	DISULFID	218	228	BY SIMILARITY.
FT	DISULFID	224	237	BY SIMILARITY.
FT	DISULFID	239	252	BY SIMILARITY.
FT	DISULFID	258	268	BY SIMILARITY.
FT	DISULFID	264	277	BY SIMILARITY.
FT	DISULFID	279	292	BY SIMILARITY.
FT	DISULFID	298	309	BY SIMILARITY.
FT	DISULFID	305	318	BY SIMILARITY.
FT	DISULFID	320	332	BY SIMILARITY.
FT	DISULFID	338	350	BY SIMILARITY.
FT	DISULFID	344	359	BY SIMILARITY.
FT	DISULFID	365	377	BY SIMILARITY.
SQ	SEQUENCE	493 AA:	54596 MW: 22DAFD70BACF1CA5 CRC64;	

Query Match	53.6%	Score 1348;	DB 1;	Length 493;
Best Local Similarity	48.0%;	Pred. No. 3.6e-93;		
Matches 235;	Conservative 76;	Mismatches 113;	Indels 66;	Gaps
Qy	17	LLLLLLSASPODEEPDSTECTDGYEWDPSDQHCRDVNECLTIPACKGEMKCINHHG	76	
Db	7	LTMTLALVKSQVTEETITTYCTDGYEWDPVQCKDIDCEDIVDPACKGKMCVNHYG	66	
Qy	77	GYLCPLRSAAVINDLHCEGPPPPVPPAAQ-----	104	
Db	67	GYLCPLKTAQII--VNNEQPQOETPAEASGSRATGTAARSMATSGVIPGGGFIASATA	124	
Qy	105	-----HPNPCPPGYEPDDQSDCVVDECAQALHDC	134	
Db	125	VAGPEVOTGRNNEVIRNPADPQRIPSNPSHRIQCAAYEQSEHNVCQDIDECTSGTHNC	184	
Qy	135	RPSDCHNLPGSYOCTCPDGYRKTPGECVIDECRY-RYCOHRCVNLPGSFRCQCPGFGQ	193	
Db	185	RLDQVCILNLKGSFYCHCLPGYKRGEGOCVIDECSPVPPYCHQGCVNTPGSEFYCQCPGFGQ	244	
Qy	194	LGPNNRSCVDVNECDMGAPCEQRCSFYGTFLCRCHOGYELHRDGFSCSDIDECSSSYL	253	
Db	245	LAANNYTCVDINECDAQNCAQOCYNILGSFICQCNGYBELSSDRLNCEDIDECRTSSYL	304	
Qy	254	CQYRCVNEPGRFSCHPQGGTQLLATRLCQDIDECESGAHOCSEAQTVCNPHFGYRCVDTN	313	
Db	305	CQYCVANEPKGFSCMCPQGYQVVRSTRTCQDINECET-TNECREDEMCWNYPHGGFCYPQN	363	
Qy	314	RCVPEYIQVSENRCCLPASNPLCRCQEPSSIVHRVMTITTSERSVPADVFOIQATSVYPGAY	373	
Db	364	PCQPPYVLTSENRCVCPVNTMCKRDVPQSIYIKYITNIRSDRSVPSDIFIQIQAITYIANTI	423	
Qy	374	NAFOIRAGNSQGDYIRIQINNFMALVARPVGTGPREYVLDLEMTNMSLMYSRASSVLR	433	
Db	424	NTFRIKSGNENGEPLRQTPSPVSAMLVLVKSLGCPREHIVGLEMLTVSSIGTERTSSVLR	483	
Qy	434	LTVEVGAYTF	443	
Db	484	LTIIVGPFSE	493	

RESULT	6
FBL5_HUMAN	
ID	FBL5_HUMAN STANDARD; PRT; 448 AA.
AC	Q9UBX5; O75966;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Fibulin-5 precursor (FIBU-5) (Developmental arteries and neural crest
DE	EGF-like protein) (Dance) (Urne p50 protein) (UP50).



RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;  
 RT "DANCE", a novel secreted RGD protein expressed in developing,  
 atherosclerotic, and balloon-injured arteries.";  
 J. Biol. Chem. 274:22476-22483(1999).  
 CC -!- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH  
 CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR  
 CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR  
 CC DEVELOPMENT AND REMODELING.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF112151; AD41767.1; -;  
 DR HSP; P00736; IAP0.  
 DR MGD; MG1:1346091; Fbln5.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR Pfam: PF00008; EGF; 4.  
 DR SMART; SM00179; EGF\_CA; 4.  
 DR SMART; SM00001; EGF\_Like; 2.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF\_2; 4.  
 DR PROSITE; PS01187; EGF\_CA; 6.  
 DR Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.  
 KW SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 1 23  
 FT DOMAIN 24 448  
 FT EGF-LIKE 1, DIVERGENT.  
 FT EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 FT EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 FT EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 FT EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 FT EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
 FT CELL ATTACHMENT SITE (POTENTIAL).  
 FT SITE 54 56  
 FT BY SIMILARITY.  
 FT DISULFID 131 144  
 FT BY SIMILARITY.  
 FT DISULFID 138 153  
 FT BY SIMILARITY.  
 FT DISULFID 155 166  
 FT BY SIMILARITY.  
 FT DISULFID 172 181  
 FT BY SIMILARITY.  
 FT DISULFID 177 190  
 FT BY SIMILARITY.  
 FT DISULFID 192 205  
 FT BY SIMILARITY.  
 FT DISULFID 211 221  
 FT BY SIMILARITY.  
 FT DISULFID 217 230  
 FT BY SIMILARITY.  
 FT DISULFID 232 245  
 FT BY SIMILARITY.  
 FT DISULFID 251 262  
 FT BY SIMILARITY.  
 FT DISULFID 258 271  
 FT BY SIMILARITY.  
 FT DISULFID 273 286  
 FT BY SIMILARITY.  
 FT DISULFID 292 305  
 FT BY SIMILARITY.  
 FT DISULFID 299 314  
 FT BY SIMILARITY.  
 FT DISULFID 320 332  
 FT BY SIMILARITY.  
 FT CARBOHYD 283  
 FT N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 296 296  
 FT N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 448 AA; 50193 MW; F15CC70CCFBFC97 CRC64;

Query Match 50.5%; Score 1269; DB 1; Length 448;  
 Best Local Similarity 48.6%; Pred. No. 2.4e-87;  
 Matches 220; Conservative 78; Mismatches 133; Indels 22; Gaps 5;  
 QY 8 LPSGLLWALLLLGASQDSEEDSYTECTDGYEWDPDSQHCSDVNECLTIPACKG 67  
 Db 1 MPGLKRLFTVITALLWLPNGNAQ-----QCTNGFDLDRQSCQCLDIDECRTIPACKG 55  
 QY 68 EMKCNHGYGLCLPRSAAVINDLHG-----EGP-----PPVPVPAQHPN-----PCPP 111  
 Db 56 DMCMVQNGGYLCIPNTNPNYRGPNPNYSTSGYPAPAAVPPASNPYPTISRLVCRF 115

QY 112 GYEPDQDSCVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYR 171  
 Db 116 GYOMDEGNQCVDDCATDSHCNPTQICINTEGGYTCSTCDGYLWLEGGQCLDIDECRYG 175  
 QY 172 YCOHRCVNLPGSPRCOCFGLGPNRSCVDVNECDMGAPCFQRCFNSVGTFLCRCHOG 231  
 Db 176 YCOQLCANPFGSYCTCNFTLLDDGRSCQDVNECTENPCVQTCVNTYVGSFICRDPG 235  
 QY 232 YELHRDGFSCSDIDEGSYSYLCQYRCVNEPGRFSCHCPQGYQLL-ATRLCQDIDECESG 290  
 Db 236 YLEEDGIIHCSMDDECSFEFLQHECVNQPGSYFCSPGVLDDNRSCQDINECHR 295  
 QY 291 AHQCSAQTCVNPFGHYRCVDTNRCVEPIYQVSENRLCPASNPCLREQSPSSVHRYMTI 350  
 Db 296 NHTCTSLQTCYNIQGGFKCIDPTSCPEEYLLIGENRCMCPAHTSCRDQPTFLYRDMDV 355  
 QY 351 TSERSVPADVFOIQATSVVPGAYNAFQIRAGNSQGFYIROIINNFMALVLRPVGTGRE 410  
 Db 356 VSGRSVPADIFQMQATRTIPGAYIIFQIKSGNREGRFYMRQTGPISATLVMTPTKIGPRD 415  
 QY 411 YVLDEMTVMNSLMYSRASSVLRLTVFVGAYTF 443  
 Db 416 IQLDLEMITVNTVINFRGSSVIRLRIYVSQYPF 448  
 RESULT 8  
 FBL5\_RAT  
 ID FBL5\_RAT STANDARD; PRT; 448 AA.  
 AC Q9WH8; O9R284;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Fibulin-5 precursor (FBL-5) (Developmental arteries and neural crest  
 DE EGF-like protein) (Dance) (Embryonic vascular EGF repeat-containing  
 DE protein) (EVEC).  
 DE FBLN5 OR DANCE.  
 GN Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99357779; PubMed=10428823;  
 RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,  
 Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,  
 Matsumori A., Sasayama S., Chien K.R., Honjo T.;  
 "DANCE", a novel secreted RGD protein expressed in developing,  
 atherosclerotic, and balloon-injured arteries.";  
 J. Biol. Chem. 274:22476-22483(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99278197; PubMed=10347091;  
 RA Kowal R.C., Richardson J.A., Milano J.M., Olson E.N.;  
 "EVEC", a novel epidermal growth factor-like repeat-containing protein  
 upregulated in embryonic and diseased adult vasculature.";  
 Circ. Res. 84:1166-1176(1999).  
 CC -!- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH  
 CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR  
 CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR  
 CC DEVELOPMENT AND REMODELING.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF112153; AD41769.1; -;



DR HSP; P07204; 1FGD.  
DR MGD; MGI:95488; Fbin2.  
DR InterPro; IPR000020; Anaphylatoxin.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR Pfam; PF01821; ANATO; 2.  
DR Pfam; PF00008; EGF; 6.  
DR SMART; SM00104; ANATO; 3.  
DR SMART; SM00179; EGF\_Ca; 9.  
DR SMART; SM00001; EGF\_Like; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 5.  
DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 3.  
DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 3.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 5.  
DR PROSITE; PS01187; EGF\_Ca; 10.  
KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;  
KW Calcium-binding; Alternative splicing; Repeat.  
FT SIGNAL 1 26  
FT CHAIN 27 1221 FIBULIN-2.  
FT DOMAIN 27 434 N.  
FT DOMAIN 27 434 SUBDOMAIN NB (CYS-RICH).  
FT DOMAIN 177 434 SUBDOMAIN NB (CYS-FREE).  
FT DOMAIN 177 434 ANAPHYLATOXIN-LIKE 1.  
FT DOMAIN 435 477 ANAPHYLATOXIN-LIKE 2.  
FT DOMAIN 478 510 ANAPHYLATOXIN-LIKE 3.  
FT DOMAIN 511 543 ANAPHYLATOXIN-LIKE 4.  
FT DOMAIN 544 594 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 595 635 EGF-LIKE 2.  
FT DOMAIN 636 669 EGF-LIKE 3.  
FT DOMAIN 670 708 EGF-LIKE 4.  
FT DOMAIN 709 755 EGF-LIKE 5.  
FT DOMAIN 756 800 EGF-LIKE 6.  
FT DOMAIN 801 846 EGF-LIKE 7.  
FT DOMAIN 847 894 EGF-LIKE 8.  
FT DOMAIN 895 937 EGF-LIKE 9.  
FT DOMAIN 938 979 EGF-LIKE 10.  
FT DOMAIN 980 1018 EGF-LIKE 11.  
FT DOMAIN 1019 1061 EGF-LIKE 12.  
FT DOMAIN 1062 1106 EGF-LIKE 13.  
FT DOMAIN 1107 1151 EGF-LIKE 14.  
FT DOMAIN 1152 1221 EGF-LIKE 15.  
FT SITE 421 423 CELL ATTACHMENT SITE (POTENTIAL).  
FT DISULFID 424 435 BY SIMILARITY.  
FT DISULFID 436 469 BY SIMILARITY.  
FT DISULFID 470 470 BY SIMILARITY.  
FT DISULFID 479 508 BY SIMILARITY.  
FT DISULFID 509 509 BY SIMILARITY.  
FT DISULFID 511 535 BY SIMILARITY.  
FT DISULFID 512 542 BY SIMILARITY.  
FT DISULFID 543 543 BY SIMILARITY.  
FT DISULFID 544 594 BY SIMILARITY.  
FT DISULFID 595 610 BY SIMILARITY.  
FT DISULFID 606 619 BY SIMILARITY.  
FT DISULFID 621 634 BY SIMILARITY.  
FT DISULFID 635 683 BY SIMILARITY.  
FT DISULFID 679 692 BY SIMILARITY.  
FT DISULFID 694 707 BY SIMILARITY.  
FT DISULFID 713 726 BY SIMILARITY.  
FT DISULFID 727 735 BY SIMILARITY.  
FT DISULFID 742 754 BY SIMILARITY.  
FT DISULFID 805 818 BY SIMILARITY.  
FT DISULFID 819 827 BY SIMILARITY.  
FT DISULFID 833 845 BY SIMILARITY.  
FT DISULFID 839 912 BY SIMILARITY.  
FT DISULFID 906 921 BY SIMILARITY.  
FT DISULFID 923 936 BY SIMILARITY.  
FT DISULFID 942 954 BY SIMILARITY.  
FT DISULFID 950 963 BY SIMILARITY.  
FT DISULFID 965 978 BY SIMILARITY.  
FT DISULFID 984 993 BY SIMILARITY.  
FT DISULFID 989 1002 BY SIMILARITY.  
FT DISULFID 1004 1017 BY SIMILARITY.  
FT DISULFID 1023 1035 BY SIMILARITY.  
FT DISULFID 1031 1044 BY SIMILARITY.  
FT DISULFID 1046 1060 BY SIMILARITY.  
FT DISULFID 1066 1079 BY SIMILARITY.

FT DISULFID 1073 1088 BY SIMILARITY.  
FT DISULFID 1093 1105 BY SIMILARITY.  
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 709 755 MISSING (IN ISOFORM EGF3-LESS).  
FT CONFLICT 140 159 HSGRYAAGHTVHLSSCRAC -> TVAVSICWYPRPLLLP  
FT CONFLICT 348 348 GF (IN REF. 2).  
FT CONFLICT 507 507 S -> L (IN REF. 2).  
FT CONFLICT 1102 1102 Q -> QQ (IN REF. 2).  
SQ SEQUENCE 1221 AA; 131818 MW; 87DB2A10A8FDC45F CRC64;  
  
Query Match 32.0%; Score 805; DB 1; Length 1221;  
Best Local Similarity 37.9%; Pred. No. 1.9e-52;  
Matches 156; Conservative 68; Mismatches 150; Indels 38; Gaps 12;  
  
QY 39 CTDGYEWDPDSQHCRDVNECLTIPACKGEMKCNHYGGLCLPRSAVINDLHGEGPPP 98  
DB 833 CMDGFLQDPEG-NCVDINECTSLLEPCRSFGSCINTVGSYTC----- 873  
  
QY 99 PVPAPQHPNCPGPGVEPDODS-CVDVDECAQALHDCRPSQDCHNLPGSVQCTCPDGYRK 157  
DB 874 ----QRNPLVCGRGTHANESEGVNCECTGVHRCGEGQLCYNLPGSYRCDCKRFGQR 929  
  
QY 158 --IGPECVDIDECRY---RYCQHRVCNLPSPRCQCEPGFQGLGNRRSCVDVNECDMGAP 212  
DB 930 DAFGRCTIDVNECWSPGRLCQHTCENTPGSYRCSAAGFLAADGKHCEVDNECET-RR 988  
  
QY 213 CQGRFNSVGTFLCRCHQGYELHRDGFSCSDIDECY-SSYLCOYRCVNEPGRFSCHCP- 270  
DB 989 CSQECANIYGSQCYCRQGYQLAEDGHTCTDECAQAGAILCTFCRVNVPGSYQACPE 1048  
  
QY 271 QGVQLLAT-RLCQDIDECESGAHQSEAOCTCVNFHGGYRCVDTNRCVPEYIOVSENRCLC 329  
DB 1049 QGYTWMANGSRCKDLDECALGTHNCEATETCHNIQGSFRCLRFD-CPPNYYRVVSQTKCER 1107  
  
QY 330 PASNPL--CREQPSIVHRYMTITTSERSYPADVFOIQTAVSYPGAYNAFOIRAGNSQGF 387  
DB 1108 TFCQDITECQTSAPRITHYQLNFTGLLVPAHFIRIGAPAFAGDTISLTITKNEEYGF 1167  
  
QY 388 YIRQINNVFAMLVLPVTPGPREYVLDLEW--TNMSLSYRASSVLRITVTF 437  
DB 1168 VTRRLNAYTGVVSLQSRVLEPRDFALDVKMLROGVSVTFLAKMYIFFTF 1219  
  
RESULT 10  
FBL2\_HUMAN STANDARD; PRT; 1184 AA.  
AC P98095;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fibrin-2 precursor.  
GN FBLN2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fibroblast;  
RX MEDLINE=95104855; PubMed=7806230;  
RA Zhang R.-Z., Pan T.-C., Zhang Z.-Y., Mattei M.-G., Timpl R.,  
RA Chu W.-L.;  
RT "Fibrin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping  
of the gene on human and mouse chromosomes.";  
RL Genomics 22:425-430(1994).  
CC -!- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS  
CC -!- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.

CC CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH  
CC CC BASEMENT MEMBRANES AND OTHER CONNECTIVE TISSUES.  
CC CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY.  
CC CC -!- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.  
CC CC -!- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.  
CC CC -----  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC CC the European Bioinformatics Institute. There are no restrictions on its  
CC CC use by non-profit institutions as long as its content is in no way  
CC CC modified and this statement is not removed. Usage by and for commercial  
CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC CC -----  
CC CC EMBL: X82494; CAA57876.1; -.  
CC CC HSP: P07204; IFGD.  
CC CC MIM: 135821; -.  
CC CC InterPro: IPR000020; Anaphylatoxin.  
CC CC InterPro: IPR000152; Asx\_hydroxyl.  
CC CC InterPro: IPR000561; EGF-like.  
CC CC InterPro: IPR001881; EGF\_Ca.  
CC CC Pfam: PF01821; ANATO; 2.  
CC CC Pfam: PF00008; EGF; 7.  
CC CC SMART: SM00104; ANATO; 3.  
CC CC SMART: SM00179; EGF\_CA; 9.  
CC CC SMART: SM00001; EGF\_like; 1.  
CC CC PROSITE: PS00010; ASX\_HYDROXYL; 5.  
CC CC PROSITE: PS01177; ANAPHYLATOXIN\_1; 3.  
CC CC PROSITE: PS01178; ANAPHYLATOXIN\_2; 3.  
CC CC PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
CC CC PROSITE: PS01186; EGF\_2; 5.  
CC CC PROSITE: PS01187; EGF\_CA; 9.  
CC CC Signal: Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;  
CC CC Calcium-binding; Repeat.  
CC CC SIGNAL 1 27 POTENTIAL.  
CC CC CHAIN 28 1184 FIBULIN-2.  
CC CC DOMAIN 28 444 N.  
CC CC DOMAIN 178 444 SUBDOMAIN NA (CYS-RICH).  
CC CC DOMAIN 445 480 SUBDOMAIN NB (CYS-FREE).  
CC CC DOMAIN 488 519 ANAPHYLATOXIN-LIKE 1.  
CC CC DOMAIN 521 553 ANAPHYLATOXIN-LIKE 2.  
CC CC DOMAIN 604 645 ANAPHYLATOXIN-LIKE 3.  
CC CC DOMAIN 679 718 EGF-LIKE 1, CALCIUM-BINDING.  
CC CC DOMAIN 719 763 EGF-LIKE 2.  
CC CC DOMAIN 764 809 EGF-LIKE 3, CALCIUM-BINDING.  
CC CC DOMAIN 810 857 EGF-LIKE 4, CALCIUM-BINDING.  
CC CC DOMAIN 858 900 EGF-LIKE 5, CALCIUM-BINDING.  
CC CC DOMAIN 901 942 EGF-LIKE 6, CALCIUM-BINDING.  
CC CC DOMAIN 943 981 EGF-LIKE 7, CALCIUM-BINDING.  
CC CC DOMAIN 1025 1069 EGF-LIKE 8, CALCIUM-BINDING.  
CC CC DOMAIN 1070 1184 EGF-LIKE 9, CALCIUM-BINDING.  
CC CC DOMAIN 1070 1184 EGF-LIKE 10, CALCIUM-BINDING.  
CC CC DOMAIN 445 472 BY SIMILARITY.  
CC CC DOMAIN 446 479 BY SIMILARITY.  
CC CC DOMAIN 459 480 BY SIMILARITY.  
CC CC DOMAIN 489 518 BY SIMILARITY.  
CC CC DOMAIN 502 519 BY SIMILARITY.  
CC CC DOMAIN 521 545 BY SIMILARITY.  
CC CC DOMAIN 522 552 BY SIMILARITY.  
CC CC DOMAIN 535 553 BY SIMILARITY.  
CC CC DOMAIN 608 620 BY SIMILARITY.  
CC CC DOMAIN 616 629 BY SIMILARITY.  
CC CC DOMAIN 631 644 BY SIMILARITY.  
CC CC DOMAIN 683 693 BY SIMILARITY.  
CC CC DOMAIN 689 702 BY SIMILARITY.  
CC CC DOMAIN 704 717 BY SIMILARITY.  
CC CC DOMAIN 723 736 BY SIMILARITY.  
CC CC DOMAIN 730 745 BY SIMILARITY.  
CC CC DOMAIN 751 762 BY SIMILARITY.  
CC CC DOMAIN 768 781 BY SIMILARITY.  
CC CC DOMAIN 775 790 BY SIMILARITY.  
CC CC DOMAIN 796 808 BY SIMILARITY.

FT DISULFID 814 827 BY SIMILARITY.  
FT DISULFID 821 836 BY SIMILARITY.  
FT DISULFID 843 856 BY SIMILARITY.  
FT DISULFID 862 875 BY SIMILARITY.  
FT DISULFID 869 884 BY SIMILARITY.  
FT DISULFID 886 899 BY SIMILARITY.  
FT DISULFID 905 917 BY SIMILARITY.  
FT DISULFID 913 926 BY SIMILARITY.  
FT DISULFID 928 941 BY SIMILARITY.  
FT DISULFID 947 956 BY SIMILARITY.  
FT DISULFID 952 965 BY SIMILARITY.  
FT DISULFID 967 980 BY SIMILARITY.  
FT DISULFID 986 998 BY SIMILARITY.  
FT DISULFID 994 1007 BY SIMILARITY.  
FT DISULFID 1009 1023 BY SIMILARITY.  
FT DISULFID 1029 1042 BY SIMILARITY.  
FT DISULFID 1036 1051 BY SIMILARITY.  
FT DISULFID 1056 1068 BY SIMILARITY.  
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1184 AA; 126543 MW; CA48490A55F9EC5D CRC64;  
  
Query Match 31.7%; Score 796.5; DB 1; Length 1184;  
Best Local Similarity 36.6%; Pred. No. 7.7e-52;  
Matches 165; Conservative 65; Mismatches 168; Indels 53; Gaps 14;  
  
QY 4 CA-----SCLPGSILLWALLLLLSASPODSEPPSYTCTDGYEWDPDSDQHCRDVNECL 59  
DB 768 CAMGHTHCQPGFL-----CONTKGSFYQARQRCMDGFLQDPG-NCVDINECT 815  
QY 60 TIPEACKGEMKCNINHYGYLCPLPSAAVINLHGEPPPPVPPAHPNCPPGYE-PDDO 118  
DB 816 SLSEPCPCPGFSCINTVGSYTC-----QRNPLICARGYHASDDG 853  
QY 119 DSCVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYR-IGPECVDIDEC---RYRYC 173  
DB 854 AKCDVYNECETGVHRCGEQGVCHNLPGSYQCDCKAFQDAGRGCIDVNECWASPGRLC 913  
QY 174 QHRCVNLPGSFRCQCPGQLGPNNRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYE 233  
DB 914 QHTCENTLGSYRCSCASGFLAADGKRCEDVNECE-AORCSQECANIYGSYCYCRQGYQ 972  
QY 234 LHRDGFSCSDIECSY-SSYLQYRCVNEPGRFSCHCP-QGYQLLAT-RLCQDIDECESG 290  
DB 973 LAEDGHTCTDIDECAGAGILCTFCLNVPGSYQACPEGGYTMANGRSCKDVEDCALG 1032  
QY 291 AHQCSAQTCVNFHGGYRCVDTNRCVPEYIQVSENRCCLCPASNPL--CREQPSIVHYRM 348  
DB 1033 THNCSEAECHNIQGSFRL-RECEPNYVQVSKTKCERTCHDFLEQCONSPARITHYQL 1091  
QY 349 TITSERVPADVFQIQATSVYPGAYNAFOIRAGNSQGDYFIROINNVFAMVLARVPTGP 408  
DB 1092 NFQTLVLPFAHIFRIGPAFTGDTIALNLIKNGSEYGTGRLNAYTGVVYLQRAVLEP 1151  
QY 409 REYVLDLEMY--TNNSIMSYRASSVLRVLFV 437  
DB 1152 RFDALDVEMKLRQGSVTTFLAKMHFFITTF 1182  
  
RESULT 11  
FBLI\_CHICK STANDARD; PRT; 684 AA.  
ID FBLI\_CHICK  
AC O73775;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fubulin-1 precursor.  
GN FBLN1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;



OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RA "Isolation of chicken and nematode fibulin-1 homologs and  
RT characterization of the nematode fibulin-1 gene";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX.  
CC -!- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF051400; AAC05388.1; -;  
DR HSP; P00742; IHCG.  
DR InterPro; IPR000020; Anaphylatoxin.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR Pfam; PF01821; ANATO; 2.  
DR Pfam; PF00008; EGF; 6.  
DR SMART; SM00104; ANATO; 3.  
DR SMART; SM00179; EGF\_Ca; 8.  
DR SMART; SM00001; EGF\_Like; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 5.  
DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 2.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS01187; EGF\_Ca; 8.  
KW Signal; Glycoprotein; Extracellular matrix; Repeat; EGF-like domain;  
KW Calcium-binding.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 684 FIBULIN-1.  
FT DOMAIN 33 74 ANAPHYLATOXIN-LIKE 1.  
FT DOMAIN 75 109 ANAPHYLATOXIN-LIKE 2.  
FT DOMAIN 110 142 ANAPHYLATOXIN-LIKE 3.  
FT DOMAIN 177 216 EGF-LIKE 1.  
FT DOMAIN 217 262 EGF-LIKE 2.  
FT DOMAIN 263 308 EGF-LIKE 3.  
FT DOMAIN 309 356 EGF-LIKE 4.  
FT DOMAIN 357 399 EGF-LIKE 5.  
FT DOMAIN 400 441 EGF-LIKE 6.  
FT DOMAIN 442 481 EGF-LIKE 7.  
FT DOMAIN 482 525 EGF-LIKE 8.  
FT DOMAIN 526 570 EGF-LIKE 9.  
FT DOMAIN 571 599 BY SIMILARITY.  
FT DISULFID 33 59 BY SIMILARITY.  
FT DISULFID 34 66 BY SIMILARITY.  
FT DISULFID 47 67 BY SIMILARITY.  
FT DISULFID 76 107 BY SIMILARITY.  
FT DISULFID 89 108 BY SIMILARITY.  
FT DISULFID 110 134 BY SIMILARITY.  
FT DISULFID 111 141 BY SIMILARITY.  
FT DISULFID 124 142 BY SIMILARITY.  
FT DISULFID 181 191 BY SIMILARITY.  
FT DISULFID 187 200 BY SIMILARITY.  
FT DISULFID 202 215 BY SIMILARITY.  
FT DISULFID 221 234 BY SIMILARITY.  
FT DISULFID 228 243 BY SIMILARITY.  
FT DISULFID 249 261 BY SIMILARITY.  
FT DISULFID 267 280 BY SIMILARITY.  
FT DISULFID 274 289 BY SIMILARITY.  
FT DISULFID 295 307 BY SIMILARITY.  
FT DISULFID 313 326 BY SIMILARITY.  
FT DISULFID 320 335 BY SIMILARITY.  
FT DISULFID 342 355 BY SIMILARITY.

FT DISULFID 361 374 BY SIMILARITY.  
FT DISULFID 368 383 BY SIMILARITY.  
FT DISULFID 385 398 BY SIMILARITY.  
FT DISULFID 404 416 BY SIMILARITY.  
FT DISULFID 412 425 BY SIMILARITY.  
FT DISULFID 427 440 BY SIMILARITY.  
FT DISULFID 446 455 BY SIMILARITY.  
FT DISULFID 451 464 BY SIMILARITY.  
FT DISULFID 466 480 BY SIMILARITY.  
FT DISULFID 486 499 BY SIMILARITY.  
FT DISULFID 495 508 BY SIMILARITY.  
FT DISULFID 510 524 BY SIMILARITY.  
FT DISULFID 530 543 BY SIMILARITY.  
FT DISULFID 537 552 BY SIMILARITY.  
FT DISULFID 557 569 BY SIMILARITY.  
FT CARBOHYD 96 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 536 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 540 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 638 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 684 AA; 75623 MW; 1638D7A094739199 CRC64;  
  
Query Match 29.7%; Score 747; DB 1; Length 684;  
Best Local Similarity 30.4%; Pred. No. 2, 1e-48;  
Matches 161; Conservative 67; Mismatches 158; Indels 144; Gaps 15;  
  
QY 23 GSASPDSEEPSYT-ECTDGYEWDPSQHQRDVNECLTIPTEACKGEMKCIHHYGYLCL 81  
Db 185 GPCSQCRDTGSSVCGSCFVQLQDPDGVNCEIDNEITGTHSCGIGTCVNTLGSFRC- 243  
QY 82 PRSAVINDLHGEPPPPVPPAQNPNCPGPEPDQDSCVDVDECAQALHDCRPSQDCH 141  
Db 244 -----QRDTSCGTGYELTDSRCKDIDECETGTHNCPPDFICQ 281  
QY 142 NLPGSYQC-----T 150  
Db 282 NTFGSRCPKPKLQCMNGFTQDALGNCIDINECLSTNMPAGQICINTDGSYTCQRISPS 341  
QY 151 CPDGY--RKIGPECVDIDEC----- 168  
Db 342 CGRGYHLNEDGTRCVDVDECSSSDQPCGEGHVINGPNRYRCKSGYSFVDSRTCDI 401  
QY 169 ----RY--RYCOHRCVNLPGSRFCQCEPGFOLGPNNRSCVDVNECDMGAPCEQRCNSVG 222  
Db 402 NECRRYPGLRCAHKCENTPGSYCTCTMGFKLSSDGRSCEDLNECE--SSPCSGECANVYG 460  
QY 223 TFLCRCHQGYELHR-DGFSQSDIDECY--SSYLQYRCVNEPGRSCHCPQ-GYQLLA- 277  
Db 461 SYQCYRRGFQLSDIDGISCEDIDECALPTGGHICFRCLINIPGSOCTCPTSGYRLAPN 520  
QY 278 TRLCQDIDECESGAHQCSBAQTCVNFPHGGYRCVDTNRCVEPIQVSENRC--LCPASNPL 335  
Db 521 ARNCQDIDECVAETHNCSFNETCFNIQGGFRCLSL-ECPEYRKRSGDTRCERLPCNENKE 579  
QY 336 CREQPSIVHYRYTITRSRVPADVFQIATSVYPCAYNAFQIRAGNSOGDFYIROINNV 395  
Db 580 COSLPLRITYYHLSPPTNTQVPTDFRMPGSPNAVPQDKILLISIQNGEGFTTKRVNHH 639  
QY 396 FAMLVLAREVTPREYVLDLEWV-----TNLSLMSYRASSVLRLTVEVGA 440  
Db 640 SGIVVMQRQITEPRDLLLTIQMLTRHGTVNTF-----IAKLFFVESA 682  
  
RESULT 12  
FBL1\_MOUSE  
ID FBL1\_MOUSE STANDARD; PRT; 705 AA.  
AC Q08879; Q08878;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fibulin-1 precursor (Basement-membrane protein 90) (BM-90).  
GN FBLN1.  
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MK31;  
 RX MEDLINE=93358997; PubMed=8354280;  
 RA Pan T.-C., Kluge M., Zhang R.Z., Mayer U., Timpl R., Chu M.-L.;  
 RA "Sequence of extracellular mouse protein BM-90/fibulin and its  
 RT calcium-dependent binding to other basement-membrane ligands.";  
 RL Eur. J. Biochem. 215:733-740(1993).  
 CC 1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX.  
 CC 1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A, B, C AND D (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER ONLY IN THEIR C-  
 CC TERMINAL REGIONS.  
 CC 1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.  
 CC 1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; X70854; CAA50207.1; -;  
 DR EMBL; X70853; CAA50206.1; -;  
 DR PIR; S36441; S36441.  
 DR HSSP; P35555; 1EMN.  
 DR MGD; MGI:95487; Fbln1.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR Pfam; PF01821; ANATO; 3.  
 DR Pfam; PF00008; EGF; 6.  
 DR SMART; SM00104; ANATO; 3.  
 DR SMART; SM00179; EGF\_Ca; 7.  
 DR SMART; SM00001; EGF-like; 2.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 3.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 3.  
 DR PROSITE; PS01186; EGF\_2; 3.  
 DR PROSITE; PS01187; EGF\_CA; 9.  
 KW Signal; Alternative splicing; Glycoprotein; Extracellular matrix;  
 KW Repeat; EGF-like domain; Calcium-binding.  
 FT SIGNAL 1 29  
 FT CHAIN 30 705  
 FT DOMAIN 36 76  
 FT DOMAIN 77 111  
 FT DOMAIN 112 144  
 FT DOMAIN 178 217  
 FT DOMAIN 218 263  
 FT DOMAIN 264 309  
 FT DOMAIN 310 357  
 FT DOMAIN 358 400  
 FT DOMAIN 401 442  
 FT DOMAIN 443 482  
 FT DOMAIN 483 526  
 FT DOMAIN 527 580  
 FT DISULFID 36 61  
 FT DISULFID 37 68  
 FT DISULFID 50 69  
 FT DISULFID 78 109  
 FT DISULFID 91 110  
 FT DISULFID 112 136  
 FT DISULFID 113 143  
 FT DISULFID 126 144  
 FT DISULFID 182 192  
 FT DISULFID 188 201  
 FT DISULFID 203 216

FT DISULFID 222 235 BY SIMILARITY.  
 FT DISULFID 229 244 BY SIMILARITY.  
 FT DISULFID 250 262 BY SIMILARITY.  
 FT DISULFID 268 281 BY SIMILARITY.  
 FT DISULFID 275 290 BY SIMILARITY.  
 FT DISULFID 296 308 BY SIMILARITY.  
 FT DISULFID 314 327 BY SIMILARITY.  
 FT DISULFID 321 336 BY SIMILARITY.  
 FT DISULFID 343 356 BY SIMILARITY.  
 FT DISULFID 362 375 BY SIMILARITY.  
 FT DISULFID 369 384 BY SIMILARITY.  
 FT DISULFID 386 399 BY SIMILARITY.  
 FT DISULFID 405 417 BY SIMILARITY.  
 FT DISULFID 413 426 BY SIMILARITY.  
 FT DISULFID 428 441 BY SIMILARITY.  
 FT DISULFID 447 456 BY SIMILARITY.  
 FT DISULFID 452 465 BY SIMILARITY.  
 FT DISULFID 467 481 BY SIMILARITY.  
 FT DISULFID 487 500 BY SIMILARITY.  
 FT DISULFID 496 509 BY SIMILARITY.  
 FT DISULFID 511 525 BY SIMILARITY.  
 FT DISULFID 531 544 BY SIMILARITY.  
 FT DISULFID 538 553 BY SIMILARITY.  
 FT DISULFID 558 579 BY SIMILARITY.  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 537 537 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 541 541 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT VARSPLIC 569 705 FQKEDTVRCIKRNDACVDPVPHVSHVLSLTPFR  
 FT EFTPEETIFRAVTPLYPANQADIFDITENLRDSDII  
 FT KRYEDMTGVVQRVIVGPFYAVIKLEMYVLGGVYSHR  
 FT NVNVHIFVSEWF -> RCARLPCHENQECPLRLITYY  
 FT HLSPTNIQVAVFVRMGPSAVPGDSMOLAITAGNERGFF  
 FT TTRKVSHSQVAULTKPIPEPDLLITVKMDLYRHGTVSSF  
 FT VAKLFIFVSAEL (IN ISOFORM D).  
 SQ SEQUENCE 705 AA; 78056 MW; FD3F06469A4BAE2B CRC64;  
 Query Match 27.8%; Score 698.5; DB 1; Length 705;  
 Best Local Similarity 35.8%; Pred. No. 8.8e-45;  
 Matches 166; Conservative 71; Mismatches 164; Indels 63; Gaps 25;  
 QY 37 TECTDGYEWDPSQHRDYNNECTIPEACKGEMKINHGYGLCLPR---SAAVINDLHG 93  
 DB 248 SSGGTGYELTEDN-NCKDIDECEETGIHNCPPDFICONTLGSFRCRPKLOCKSGFIQDALG 306  
 QY 94 E-----GPPPPVPPAQH-----PNPCPPGYEPDODS-CYDVDSACA 130  
 DB 307 NCIDINECLISAPCPVGTQNTGEGSYTCQKNVNP-CGRGYHLNEEGRCVDVDECAFP 365  
 QY 131 LHDRCPSQDCHNLPGSYQCTPDG--YRKIGPECVDDIDEC-RY--RYCOHRCVNLPGSFR 185  
 DB 366 AEPCKGKHCLNSPGRCECKAGFYFDGISTRCTVDINECQRYPGRLCGHKCENTPGSEH 425  
 QY 186 CQCEPGFQLGPNNRSCVDYNECDMGAPCQRNSYGTFLCRCHQGYELHR-DGFGSCDI 244  
 DB 426 CSCSAGFRLSVDRGSCEDVNEC-LNSPSCQECANVYGSYQCYCRRGYQLSDVDGVTCEDI 484  
 QY 245 DECSY--SSYLCOYRCVNEBGRFSCHCP-QGYQLLAT-RLCDDIDECESGASQEAQTC 300  
 DB 485 DECALPTGTHICSYRCINIPGSPQCSGGRYFLAPNGRNCQDIDECVGIHNCISNEYC 544  
 QY 301 VNFHGGYRCVDTNRCVPEYITQVSEN-----RCL--CPASNPLC-REQPSSIVHRYM 348  
 DB 545 FNIQGSFRCL-SPECENYRRSADTFQEKTDIVRCIKSRPNDEACVDPVHTVSHTVI 603  
 QY 349 TITTSERSV--PADVFOIQA-TSVYPG--AYNAFOIIRAGNSQGF-VIRQINNFMAMLV- 401  
 DB 604 SLPTFRETRPEETIFRAVTPLYPANQADIFDITENLRDSDIFIKRYEDGMTGVVVR 663  
 QY 402 -ARPVTGPREYVLDLEM-VTMNSLSMYRASSYRLRLTFVEGAYTF 443  
 DB 664 QVRPIVGPFFAVUKLENNYVLGGVYSHR--NVNVHIFVSEWF 705



Query Match 26.4%; Score 664.5; DB 1; Length 703;  
Best Local Similarity 36.5%; Pred. No. 3e-42;  
Matches 160; Conservative 62; Mismatches 157; Indels 59; Gaps 24;

QY 38 ECTDGYEWDPSOHRDVECLTIPACKGEMKCNHYGCLYCLPRSAVINDLHGEGPP 97  
Db 293 QCKSGFTQDA-LGNCIDINECLISAPCPGHTCINTEGSYTC----- 334

QY 98 PPVPPAQHPNCPGPEYEPDODS-CYDVDECAQALHDCRPSQDCHNLPGSYOCTCPDGY- 155  
Db 335 ----QKNVFN-CGRHNEEGTRCVDDVDECAPPAEPCKGHRVNSPGRFCECKTGY 389

QY 156 -RKIGPECVDIDRC-RY-RYOHRQCNVLPGRFCOCEPFGQLGPNRNSCDVDNCEMDGA 211  
Db 390 FDGISMVCDVNECQYRPGRLCGHKCENTLGSVLCSCSVGFRSLVSDGRSCEDINECS-SS 448

QY 212 PCORCFNSVGTFLCRCHQGYELHR-DGFSKSDIDRCSY--SSYLQYRCVNEPGRFSCH 268  
Db 449 CSQECANVYGYOCYRCRGYQLSDVGVTCEDIDECALPTGGHICSYRCINIPGSFQCS 508

QY 269 CP-QGYQLLAT-RLCQIDBCEGSAHQCSQAOTCVNFHGGYRCVDTNRCVEPY-----I 320  
Db 509 CPSSGYELAPNGNCQIDECVTHGNCINETCFNLQGAFCRL-AFECPENVYRRAATL 567

QY 321 QVSEN---RCL--CPASNPICREOP-SSIVHYWTITTSERSV--PADVFOIQH-TSVYPG 371  
Db 568 QBEKTDIVRCIKSCRPNVTCVDPVHTTISHTVLSLTFRETRPEETIFLRAITPPHPA 627

QY 372 --AYNAFOIRAGNSQGFYI--QOINVPFAMLY-LARPVTGPREYVLDLDM-VTMNSLMS 425  
Db 628 SQANIPDITEGLNRSDFDIKRYMDGTVGVVVRQRPVIGPPHVALKLEMMYVGVGVS 687

QY 426 YRASSRLRLVFFVGAYTF 443  
Db 688 HR--NVNVNRFVSEYWF 703

RESULT 14  
FBL1\_CAEEL STANDARD; PRT; 712 AA.  
AC 077469; 077474;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fibulin-1 precursor.  
GN FBLN1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;  
RT "Isolation of chicken and nematode fibulin-1 homologs and  
characterization of the nematode fibulin-1 gene."  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX (BY SIMILARITY).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; C (SHOWN HERE) AND D; ARE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF051403; AAC28323.1; -.  
DR EMBL; AF051403; AAC28324.1; -.  
DR EMBL; AF051403; AAC28322.1; -.  
DR HSSP; P00736; IAPQ.  
DR InterPro; IPR000020; Anaphylatoxin.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR Pfam; PF01821; ANATO; 2.  
DR Pfam; PF00008; EGF; 6.  
DR SMART; SM00104; ANATO; 2.  
DR SMART; SM00179; EGF\_Ca; 6.  
DR SMART; SM00001; EGF-like; 3.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
DR PROSITE; PS01186; EGF-2; 5.  
DR PROSITE; PS01187; EGF\_Ca; 8.  
KW Signal; Glycoprotein; Extracellular matrix; Repeat; EGF-like domain;  
KW Calcium-binding; Alternative splicing.  
FT SIGNAL 1 17  
FT CHAIN 18 712  
FT FIBULIN-1.  
FT DOMAIN 23 64  
FT DOMAIN 65 96  
FT DOMAIN 97 129  
FT DOMAIN 155 194  
FT DOMAIN 195 241  
FT DOMAIN 242 303  
FT DOMAIN 304 350  
FT DOMAIN 351 390  
FT DOMAIN 391 434  
FT DOMAIN 435 475  
FT DOMAIN 476 520  
FT DOMAIN 521 568  
FT DISULFID 23 49  
FT DISULFID 24 56  
FT DISULFID 37 57  
FT DISULFID 66 94  
FT DISULFID 79 95  
FT DISULFID 97 121  
FT DISULFID 98 128  
FT DISULFID 111 129  
FT DISULFID 159 168  
FT DISULFID 164 178  
FT DISULFID 180 193  
FT DISULFID 199 212  
FT DISULFID 206 221  
FT DISULFID 227 240  
FT DISULFID 246 259  
FT DISULFID 253 268  
FT DISULFID 291 302  
FT DISULFID 308 320  
FT DISULFID 314 329  
FT DISULFID 336 349  
FT DISULFID 355 365  
FT DISULFID 360 374  
FT DISULFID 376 389  
FT DISULFID 395 409  
FT DISULFID 403 418  
FT DISULFID 420 433  
FT DISULFID 439 450  
FT DISULFID 446 459  
FT DISULFID 461 474  
FT DISULFID 480 495  
FT DISULFID 491 504  
FT DISULFID 506 519  
FT DISULFID 525 537  
FT DISULFID 530 546  
FT DISULFID 551 565  
FT DISULFID 551 565  
FT VARSPLIC 564 712  
RCNROPSACGLPEECISKVPLFLTYQFISLARAVPISSHRPA  
ITLFKVSAPNADTENVFELQKLTIVGAPNVLPAINFL  
LQKGRNSAVVTLDSLDGQVTKLQLLRMSKKNFT  
YARNLIVVAANKRINTVHPPLMKIR -> QIADGYSCKV  
CSTDETECLNGREVLQFRAVPVSKLTISPIEVSRIYTH  
MGVPFSDVNDLVGQRHFRIVYQERNIGIVQLVKPISGPTV

```
FT          ETKVNIHTKSRGTGVLAFNAFNEALIEISVSKYFF (IN
SQ          ISOFORM D).
SEQUENCE   712 AA; 77009 MW; 52CE8CF8BF296BC5 CRC64;

Query Match      21.9%; Score 549.5; DB 1; Length 712;
Best Local Similarity 26.5%; Pred. No. 1e-33;
Matches 141; Conservative 61; Mismatches 137; Indels 193; Gaps 20;

QY 39 CTGDEWDPDSQHRDQVNECLTIPEACKGEMKINHYGGYLCLPRSAVINDLHGEGPP 98
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 180 CRSGFDLAPDGMACVDDICATLMDCLLESQRCNLTPGSKFI-RTLS----- 226

QY 99 PVPPAHPNCPGPGYEDDQ-DSVDVDECAQALHDCRSQDCHNLPGSYQC----- 149
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 227 -----CGTGYAMDSETERCDVDECNLGSBDCGLYQCRNTQGSYRCDAKRCGDG 276

QY 150 -----TCPDGYRKIGPEGVDIDEC-----RVRVC 173
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 277 ELONPMTGECTSITCPNGYYPKMGKNDIDECVTGHCAGAGECVNTPGSRQOQGNLC 336

QY 174 QH-----RCVNLP----- 181
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 337 AHGYEVNGATGFCEDYNECOQGVCGSMECINLPCTYKCKGPGYEFNDAKKRCEDVDECI 396

QY 182 -----GSFRCEGFGFQLPNNRSCVDVNECDMG-APCEQRCFNSYGT 223
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 397 KFAHVCYDLASABCTINTGTSFECKKPGFQLASDGRRCEDVNECTTGIAACEQRCVNIPGS 456

QY 224 FLCRCHQGVHLDGFGSCSDSCSY-----SSYLQYRCVNEPGRSFCHPCPGYQLLAT- 278
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 457 YQICDRGFALGPDGTYKCEDIDECSTWAGSGNDLGMGGCINTKGSYLQCPGPKTOPDG 516

QY 279 RLQODIDECEGSAHQCSA-QTCVNFHGGYRCVDTNRCVEPIYQVSENRCLC---PASNP 334
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 517 RTCVDVDECAMG--ECAGSKVKCVNTLGSFKCHSID-CPTNYTHDSLNNKRNCRQPSACG 573

QY 335 L---CREQPSIVHRYWTITSERSVP-----ADV---FQIQATSVYP 370
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 574 LPCECKVPLFLYQYFISLA--RAVPISSHRPATITFKVSAPNHADTEYVNFELQKTTIV 631

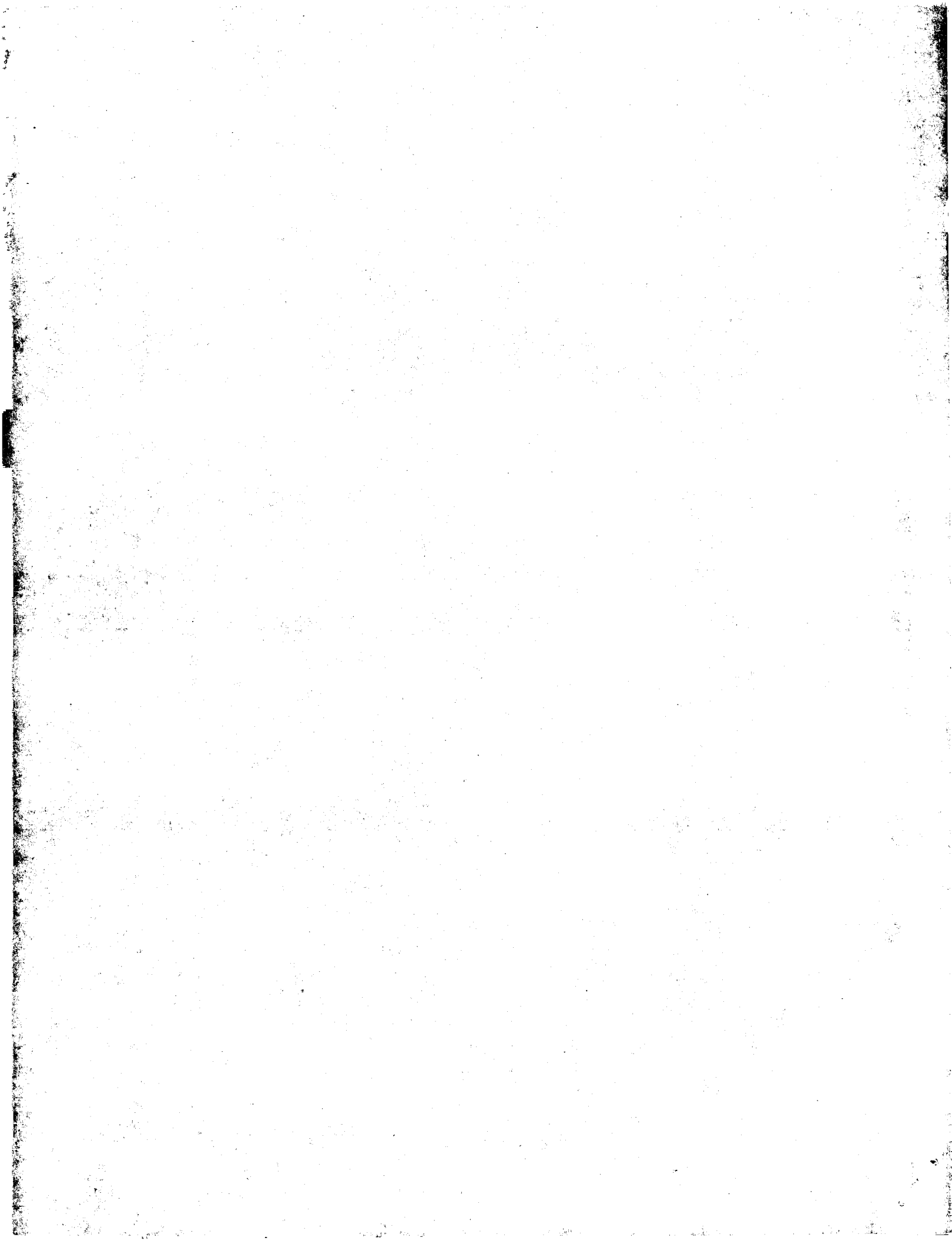
QY 371 GAYNA-----FQIRAGNSQSGFYIRQINNVFAMLVLPARVPTGREYVLDL 415
   | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 632 GAFNVLPAIRANFLLOKGEKNS-----AVVTLRSLDGLPGTQVKLQL 673

RESULT 15
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguineti C.,
RA Bonadio J., Mecham R.P., Ramirez F.;
RT "Structure and expression of fibrillin-2, a novel microfibrillar
RT component preferentially located in elastic matrices.";
RL J. Cell Biol. 124:855-863(1994).
RN [2]
RP SEQUENCE OF 752-1505 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
RA Tsiouras P., Ramirez F., Hollister D.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
```

```
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN [3]
RX VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
RP MEDLINE=96083599; PubMed=7493032;
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
RT congenital contractural arachnodactyly.";
RL Nat. Genet. 11:456-458(1995).
RN [4]
RX VARIANTS CCA HIS-1114.
RP MEDLINE=98407789; PubMed=9737771;
RA Babcock D., Gasner C., Francke U., Maslen C.;
RT "A single mutation that results in an asp-to-his substitution and
RT partial exon skipping in a family with congenital contractural
RT arachnodactyly.";
RL Hum. Genet. 103:22-28(1998).
RN [5]
RX VARIANTS CCA PHE-1141 AND TRP-1252.
RP MEDLINE=20259336; PubMed=10797416;
RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,
RA Godfrey M.;
RT "Two novel fibrillin-2 mutations in congenital contractural
RT arachnodactyly.";
RL Am. J. Med. Genet. 92:7-12(2000).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
CC ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT AFFECT THE
CC AORTA AND THE EYES.
CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: U03272; AA18950.1; -
CC EMBL: X62009; -; NOT_ANNOTATED_CDS.
CC PIR: S17063; S17063.
CC PIR: S31101; S31101.
CC HSSP: P35555; 1EMN.
CC MIM: 121050; -
CC InterPro: IPR000152; Asx_hydroxyl.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001881; EGF_Ca.
CC InterPro: IPR001438; EGF_II.
CC InterPro: IPR002212; TB.
CC Pfam: PF00008; EGF; 45.
CC Pfam: PF00683; TB; 9.
CC PRINTS: PR00010; EGFBL00D.
CC SMART: SM00179; EGF_CA; 43.
CC SMART: SM00001; EGF_like; 3.
CC PROSITE: PS00010; ASX_HYDROXYL; 43.
CC PROSITE: PS00022; EGF_1; 2.
CC PROSITE: PS01186; EGF_2; 37.
CC PROSITE: PS01187; EGF_CA; 43.
CC Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2911 FIBRILLIN 2.
FT DOMAIN 111 142 EGF-LIKE 1, NON-CALCIUM BINDING.
FT DOMAIN 145 176 EGF-LIKE 2, NON-CALCIUM BINDING.
FT DOMAIN 176 207 EGF-LIKE 3, NON-CALCIUM BINDING.
FT DOMAIN 275 316 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 317 358 EGF-LIKE 5, CALCIUM-BINDING.
```



Job time: 249 sec





GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2002, 09:42:15 ; Search time 31.1 Seconds

(without alignments)  
2464.204 Million cell updates/sec

Title: US-09-829-936A-22

Perfect score: 2513

Sequence: 1 MLPACSLPGSLLLWALLLL.....MSYRASSVLRLTVFGAYTF 443

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rhodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvrius:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2501	99.5	443	4	Q96TF5	Q96tf5 homo sapien
2	2473	98.4	443	4	Q9H3D5	Q9h3d5 homo sapien
3	2397	95.4	443	11	Q9JM06	Q9jm06 mus musculu
4	808	32.2	1174	11	Q99K58	Q99k58 mus musculu
5	797.5	31.7	576	4	Q9F3V7	Q9f3v7 homo sapien
6	744	29.6	685	11	Q922K8	Q922k8 mus musculu
7	686.5	27.3	704	13	Q73774	O73774 gallus gall
8	653.5	26.0	681	13	Q42182	Q42182 brachydanio
9	649	25.8	495	4	Q9HBQ5	Q9hbq5 homo sapien
10	560	22.3	554	4	Q9UH16	Q9uh16 homo sapien
11	555.5	22.1	689	5	Q95N23	Q95n23 caenorhabdi
12	555	22.1	589	5	Q9T2S1	Q9t2s1 caenorhabdi
13	529.5	21.1	798	5	O18026	O18026 caenorhabdi
14	523.5	20.8	1833	11	O08999	O08999 mus musculu
15	523	20.8	2906	11	Q9WH9	Q9wh9 rattus norv
16	522	20.8	2809	4	Q96JP8	Q96jp8 homo sapien

## ALIGNMENTS

RESULT 1

Q96TF5 PRELIMINARY; PRT; 443 AA.

AC Q96TF5; 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE MUTANT P53 BINDING PROTEIN 1 (MBP1).  
 GN MBP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tanaka S.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Tanaka S., Sugimachi K., Sugimachi K.;  
 RT "Human mutant p53 binding protein (MBP1).";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB030655; BAA92880.1;  
 SQ SEQUENCE 443 AA; 49421 MW; 9CE175F4F388A56D CRC64;

Query Match 99.5%; Score 2501; DB 4; Length 443;  
 Best Local Similarity 99.5%; Pred. No. 1.le-246;  
 Matches 441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLPACSLPGSLLLWALLLLLSASPQSDSYTECTDGYEWDPSQHQCRVNECLT 60

Db 1 MLPCTSLPGSLLLWALLLLLSASPQSDSYTECTDGYEWDPSQHQCRVNECLT 60

QY 61 IPEACKGEMKCINHYGYLCLPRSAVINLHGEPPPPVPPPAQHPNCPGYPEPDQDS 120

Db 61 IPEACKGEMKCINHYGYLCLPRSAVINLHGEPPPPVPPPAQHPNCPGYPEPDQDS 120

QY 121 CVDVDECAALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRCQHCRCVNL 180

O35806 rattus norv  
 Q28019 bos taurus  
 Q96hb9 homo sapien  
 Q9ns15 homo sapien  
 Q9h7k2 homo sapien  
 Q9v4b8 drosophila  
 Q9wh8 rattus norv  
 Q14767 homo sapien  
 Q9vs89 drosophila  
 Q88840 mus musculu  
 Q9blj1 ciona intes  
 O75412 homo sapien  
 O00508 homo sapien  
 Q96ec7 homo sapien  
 Q96rwt homo sapien  
 Q96k89 homo sapien  
 P87363 gallus gall  
 Q9np01 homo sapien  
 Q88349 mus musculu  
 Q60784 mus musculu  
 Q61810 mus musculu  
 Q25678 podocoryne  
 Q9tvq2 caenorhabdi  
 Q60789 mus musculu  
 Q9vt08 drosophila  
 Q9cxd8 mus musculu  
 Q9daw5 mus musculu  
 Q88281 rattus norv  
 Q9bi05 eimeria ten

```

|||||
Db 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRQCQRCVNL 180
QY 181 PGSFRCQCEPGFQGLGNNSRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240
Db 181 PGSFRCQCEPGFQGLGNNSRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240
QY 241 CSDIDECSSYLCOYRCVNEPGRFSCHCPQGYQLLATRLCQDIDECESGAHQCEAQC 300
Db 241 CSDIDECSSYLCOYRCVNEPGRFSCHCPQGYQLLATRLCQDIDECESGAHQCEAQC 300
QY 301 VNFHGGYRCVDTNRCVPEYIOVSENRCCLCPASNPLCREQPSIVHRYMTITTSERSVPADV 360
Db 301 VNFHGGYRCVDTNRCVPEYIOVSENRCCLCPASNPLCREQPSIVHRYMTITTSERSVPADV 360
QY 361 FQIQATSVTPGAYNAFOIRAGNSQGFYIRQINNVSAMLVLARPVTPGREYVLDLEMTM 420
Db 361 FQIQATSVTPGAYNAFOIRAGNSQGFYIRQINNVSAMLVLARPVTPGREYVLDLEMTM 420
QY 421 NSLSYRASSVLRLTVFVGAYTF 443
Db 421 NSLSYRASSVLRLTVFVGAYTF 443

```

```

RESULT 2
ID Q9H3D5 PRELIMINARY; PRT; 443 AA.
AC Q9H3D5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FIBULIN-LIKE EXTRACELLULAR MATRIX PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seibold S., Marx M.;
RT "Cloning of a new fibulin-like gene.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF124486; AAG45245.1; -.
DR HSSP; P35555; 1EMN.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 3.
DR PRINTS; PR00907; THRBOMODULN.
DR SMART; SM00181; EGF; 5.
DR SMART; SM00179; EGF_Ca; 6.
DR SMART; SM00001; EGF_like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR EGF-like domain; Glycoprotein; Hydroxylation; Matrix protein.
KW EGF-like domain; Glycoprotein; Hydroxylation; Matrix protein.
SQ SEQUENCE 443 AA; 49535 MW; D91784BF36A8A060 CRC64;

```

```

Query Match 98.4%; Score 2473; DB 4; Length 443;
Best Local Similarity 98.9%; Pred. No. 8e-244;
Matches 438; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLPASCCLPGSLLLWALLLLLLGASPDSEPDSTECTDGYEWDPSQCHRDVNECLT 60
Db 1 MLPASCCLPGSLLLWALLLLLLGASPDSEPDSTECTDGYEWDPSQCHRDVNECLT 60
QY 61 IPEACKGEMKCINHYGYLCIPRASAIVNDLHGEGPPPPVPPAHPNPPCPGYPDDQDS 120
Db 61 IPEACKGEMKCINHYGYLCIPRASAIVNDLHGEGPPPPVPPAHPNPPCPGYPDDQDS 120
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRQCQRCVNL 180
Db 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRQCQRCVNL 180

```

```

QY 181 PGSFRCQCEPGFQGLGNNSRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240
Db 181 PGSFRCQCEPGFQGLGNNSRSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240
QY 241 CSDIDECSSYLCOYRCVNEPGRFSCHCPQGYQLLATRLCQDIDECESGAHQCEAQC 300
Db 241 CSDIDECSSYLCOYRCVNEPGRFSCHCPQGYQLLATRLCQDIDECESGAHQCEAQC 300
QY 301 VNFHGGYRCVDTNRCVPEYIOVSENRCCLCPASNPLCREQPSIVHRYMTITTSERSVPADV 360
Db 301 VNFHGGYRCVDTNRCVPEYIOVSENRCCLCPASNPLCREQPSIVHRYMTITTSERSVPADV 360
QY 361 FQIQATSVTPGAYNAFOIRAGNSQGFYIRQINNVSAMLVLARPVTPGREYVLDLEMTM 420
Db 361 FQIQATSVTPGAYNAFOIRAGNSQGFYIRQINNVSAMLVLARPVTPGREYVLDLEMTM 420
QY 421 NSLSYRASSVLRLTVFVGAYTF 443
Db 421 NSLSYRASSVLRLTVFVGAYTF 443

RESULT 3
ID Q9JM06 PRELIMINARY; PRT; 443 AA.
AC Q9JM06;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EGF-CONTAINING FIBULIN-LIKE EXTRACELLULAR MATRIX PROTEIN 2.
GN EFEMP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Katsanis N., Venable S., Smith J.R., Lupski J.R.;
RT "Isolation of a paralog of the Doyme honeycomb retinal dystrophy gene
RT from the multiple retinopathy critical region on 11q13.";
RL Hum. Genet. 106:66-72(2000).
DR EMBL; AF109122; AAF65189.1; -.
DR HSSP; P00736; IAPQ.
DR MGD; MGI:1891209; Efemp2.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001491; Thrbomodulin.
DR PRINTS; PR00907; THRBOMODULN.
DR SMART; SM00179; EGF_Ca; 4.
DR SMART; SM00001; EGF_like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_Ca; 6.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Matrix protein; Repeat.
SQ SEQUENCE 443 AA; 49452 MW; 5AEC2A91048B336A CRC64;

```

```

Query Match 95.4%; Score 2397; DB 11; Length 443;
Best Local Similarity 95.0%; Pred. No. 4.5e-236;
Matches 421; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MLPASCCLPGSLLLWALLLLLLGASPDSEPDSTECTDGYEWDPSQCHRDVNECLT 60
Db 1 MLPASCCLPGSLLLWALLLLLLGASPDSEPDSTECTDGYEWDPSQCHRDVNECLT 60
QY 61 IPEACKGEMKCINHYGYLCIPRASAIVNDLHGEGPPPPVPPAHPNPPCPGYPDDQDS 120
Db 61 IPEACKGEMKCINHYGYLCIPRASAIVNDLHGEGPPPPVPPAHPNPPCPGYPDDQDS 120
QY 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRQCQRCVNL 180
Db 121 CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRQCQRCVNL 180

```

Db 121 CVDVDECTQALHDCRPSQDCHNLPGSYQCTCPDGYRKIGPECVDIDECRYRYCOHRCVNL 180  
 QY 181 PGSFRCQCEPGFOLGPNNSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240  
 Db 181 PGSFRCQCEPGFOLGPNNSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHRDGF 240  
 QY 241 CSDDIDECSSYLQYRCVNEPGRFSCCHCPQGVQYLLATRLCQDIDECESGAHQCSEAQC 300  
 Db 241 CSDDIDECSSYLQYRCVNEPGRFSCCHCPQGVQYLLATRLCQDIDECESGAHQCSEAQC 300  
 QY 301 VNFHGYRCVDTNRCVPEYQVSENRCCLPASNPCLREQSPSSIVHRYMTTTSRSPADV 360  
 Db 301 VNFHGYRCVDTNRCVPEYQVSDNRCCLPASNPCLREQSPSSIVHRYMTTTSRSPADV 360  
 QY 361 FQIQATSVYPGAYNAFQIRAGNSQGDYFIQINNVFAMVLARVPTGPREYVLDLEVM 420  
 Db 361 FQIQATSVYPGAYNAFQIRSGNTQGDYFIQINNVFAMVLARVPTGPREYVLDLEVM 420  
 QY 421 NSLMSYRASSVLRLTVFGAYTF 443  
 Db 421 NSLMSYRASSVLRLTVFGAYTF 443

RESULT 4  
 Q99K58 PRELIMINARY; PRT; 1174 AA.  
 AC Q99K58; ID Q99K58; DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE SIMILAR TO FIBULIN 2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=MAMMARY TUMOR. WAP-TGF ALPHA MODEL. 7 MONTHS OLD, GROSS  
 RC TISSUE.;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC005443; AAH05443.1; -  
 DR HSSP; P00736; IAPQ.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR Pfam; PF01821; ANATO; 2.  
 DR Pfam; PF00008; EGF; 6.  
 DR SMART; SM00104; ANATO; 3.  
 DR SMART; SM00181; EGF; 11.  
 DR SMART; SM00179; EGF\_CA; 9.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 3.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 3.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 5.  
 DR PROSITE; PS01186; EGF 2; 5.  
 DR PROSITE; PS01187; EGF\_CA; 9.  
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 FT NON\_TER 1  
 SQ SEQUENCE 1174 AA; 126460 MW; 8d628AC710FBA6B8 CRC64;

Query Match 32.2%; Score 808; DB 11; Length 1174;  
 Best Local Similarity 38.1%; Pred. No. 1.9e-73;  
 Matches 157; Conservative 67; Mismatches 150; Indels 38; Gaps 12;  
 QY 39 CTDGYEWDPSQHCVRDNECLTIPEACKGMKCNHGYGLCLPRSAAVINDLHGEGPPP 98  
 Db 786 CMDGFLQDPEG-NCVDINETSLLPCRGFCINTVGSYTC----- 826  
 QY 99 PVPAPQHNPCCPGYEPDDDS-CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRK 157  
 Db 827 ----QRNPLVCGRGYHANEESGECVDVNECETGVHRCGEGQLCYNLPGSYRCDCKPGFOR 882

QY 158 --IGPECVDIDECRY---RYCOHRCVNLPGSFRCQCEPGFOLGPNNSCVDVNECDMGAP 212  
 Db 883 DAFGRCTIDVNECWSPGRLCQHTCCTGYSYRSCAAGFLAADKGHCEDVNECET-RR 941  
 QY 213 CEQRCFNSYGTFLCRCHQGYELHRDGFSCSDIDECYSY-SSYLQYRCVNEPGRFSCCHCP- 270  
 Db 942 CSQECANIYGSQYCYCRQGYQLAEDGHTCTDIDECAGAGILCTFCRVNVPGSYQACPE 1001  
 QY 271 QGYQLLAT-RLCQDIDECESGAHQCSEAQTCVNFHGYRCVDTNRCVPEYQVSENRCCLC 329  
 Db 1002 QGYTMANGRSCKDLDECALGTHNCSEAECHNIQGSFRCFLRED-CPNPNVRYSETKCR 1060  
 QY 330 PASNPL-CREQSPSSIVHRYMTTTSRSPADVFOQATSVYPGAYNAFQIRAGNSQGD 387  
 Db 1061 TTCQDITEQTSPTARITHYQLNFTQGLLPVAFHIFRIGPAPAFAGDTISLTITKNGE 1120  
 QY 388 YTRQINNVFAMVLARVPTGPREYVLDLEMV--TMSLMSYRASSVLRLTVF 437  
 Db 1121 VTRRLNAYTGVSQSLQSVLEPRDFALDVENMKLRQGSVTFLLAKMYIFFTTF 1172  
 RESULT 5  
 Q9Y3V7 PRELIMINARY; PRT; 576 AA.  
 AC Q9Y3V7; ID Q9Y3V7; DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYPOTHETICAL 63.3 KDA PROTEIN (FRAGMENT).  
 GN DKF2P586A1519.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=UTERUS;  
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL050095; CAB43267.1; -  
 DR HSSP; P00736; IAPQ.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_CA.  
 DR Pfam; PF00008; EGF; 6.  
 DR SMART; SM00179; EGF\_CA; 8.  
 DR SMART; SM00001; EGF\_LIKE; 2.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
 DR PROSITE; PS01186; EGF 2; 4.  
 DR PROSITE; PS01187; EGF\_CA; 9.  
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;  
 KW Hypothetical protein; Repeat.  
 FT NON\_TER 1  
 SQ SEQUENCE 576 AA; 63274 MW; FF2F0E9B185D8AC3 CRC64;

Query Match 31.7%; Score 797.5; DB 4; Length 576;  
 Best Local Similarity 36.6%; Pred. No. 9.3e-73;  
 Matches 165; Conservative 65; Mismatches 168; Indels 53; Gaps 14;

QY 4 CA---SCLPGSILLWALLLLLGSSASPODSEEPDSEYTCCTGCEYWDPSQHCVRDNECL 59  
 Db 160 CAMGTTTCQGFJ-----CONTRGFSYCARQAROMDGLQDPEG-NCVDINECT 207  
 QY 60 TIPEACKGMKCNHGYGLCLPRSAAVINDLHGEGPPPVPVPAQHNPCCPGYE-PDQ 118  
 Db 208 SLSEPCRGFCINTVGSYTC-----QRNPLICARGYHASDDG 245  
 QY 119 DSCVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGYRK--IGPECVDIDEC---RYRYC 173  
 Db 246 TKCDVDNECETGVHRCGEGQVCHNLPGSYRCDCKAGFQDAFGRGICIDVNECWASPGRLC 305  
 QY 174 QHRCVNLPGSFRCQCEPGFOLGPNNSCVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYE 233

```

Db 306 QHTCENTLGSYSCASGELLAADKRCEDVNECE-AQRCSECANIYGYQCYCROGYQ 364
QY 234 LHRDGFSCSDIDECY-SSYLQYRCVNEPGRFSCHP-QGYQLLAT-RLCODIDECESG 290
Db 365 LAEDGHTCTDIDECAGAGILCTFCLNVPGSYQACPEQGYTMTANGRSCKDVECALG 424
QY 291 AHOCSEAOQCVNPHGRCVDRNRCVEPIQVSENRCILCPASNPL--CREQPSIVHRYM 348
Db 425 THNCSAEFTCHNIQGSFRL-RFECPPNVQVSKTKCERTTCHDFLEQNSPARITHYOL 483
QY 349 TITSESVPADVFOIQATSVYFAYNAFOIRAGNSOGDFYIRQINNFMVLVLRPVGCP 408
Db 484 NFQGTGLVPAHFRIGRAPAFATGDTIALNIKGNBEGYGTRELRNAYTCGVVQLQRAVLBP 543
QY 409 REVVLDEMV--TMNSLMSYRASSVLRLTVF 437
Db 544 RFDALDEMKLWRQGSVTTFLAKMHFFTTF 574

RESULT 6
ID Q922K8 PRELIMINARY; PRT; 685 AA.
AC Q922K8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE SIMILAR TO FIBULIN 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007140; AAH07140.1; -.
SQ SEQUENCE 685 AA; 75283 MW; EF0D77D7F66B73B8 CRC64;

Query Match 29.6%; Score 744; DB 11; Length 685;
Best Local Similarity 37.0%; Pred. No. 3.3e-67;
Matches 165; Conservative 66; Mismatches 163; Indels 52; Gaps 18;

QY 37 TECTDGYEWDPSQCHRDVNECLTIPACKGEMKCINHYGGYLCPLPR---SAAVINDLAG 93
Db 248 SSCGTGYELTEDN-NCKDIDECETGIHNCPPDFICQNTLGSFRCRPLKQCKSGFIQDALG 306
QY 94 E-----GPPPPVPPAQH-----PNCPPGYEPDDODS-CVDVDECAQA 130
Db 307 NCIDINECLTSAPCPYVGQTCINTEGYSYTCQKNVNP-CGRGYHLNEEGTRCVDVDECSPP 365
QY 131 LHDCRPQDCHNLPGSYQCCPDG--YRKIGPECVDIDEC-RY--RYCOHRCVNLPGSFR 185
Db 366 AEPCKGHCLNPSGRCCKAGFYDGLSRITCVLINEQRPGRGLCHKCENTGSGFH 425
QY 186 QCCEPFGQLGNPNRNSVDVNECDMGAPCEQRCFNSYGTFLCRCHQGYELHR-DGFCSDI 244
Db 426 CSCSAGFRLSVDRSCEDVNEC-LNSPCSECANVYGYCYCRRGYQLSDVDGVTCEDI 484
QY 245 DECSY--SSYLQYRCVNEPGRFSCHP-QGYQLLAT-RLCODIDECSEAHOCSEAQTC 300
Db 485 DECALPTGGHTICSYRCINIFGSCFPCGYRLAPNRCNODIDECVTGIHNCINETC 544
QY 301 VNEHGGYRCVDTNRCVEPIQVSENRC--LCPASNPLCREQPSIVHRYMTITSESVPA 358
Db 545 FNIQGSFRL--SFCEPENYRSADTRCERLPCHENQECRPLRLITYHLSFPTNIQVPA 603
QY 359 DVFOIQATSVYFAYNAFOIRAGNSOGDFYIRQINNFMVLVLRPVGTPREYVLDLEMY 418
Db 604 VFERMGSSAVPGDSMLAITAGNEEGFTTRKVSHRHSGVVALTKP IPEPRDLLLTVMKD 663
QY 419 TMNSLMSYR----ASSVLRLTVFVGA 440
```

```

Db 664 L-----YRHGTVSSFVAKLFIYVSA 683

RESULT 7
ID O73774 PRELIMINARY; PRT; 704 AA.
AC O73774;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE FIBULIN-1, ISOFORM D.
GN FBLN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99120531; PubMed-9923656;
RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;
RT "Identification of chicken and C. elegans fibulin-1 homologs and
characterization of the C. elegans fibulin-1 gene.";
RL Matrix Biol. 17:635-646(1998).
DR EMBL; AF051399; AAC05387.1; -.
DR HSSP; P00742; IHCG.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 2.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00104; ANATO; 3.
DR SMART; SM00179; EGF_CA; 8.
DR SMART; SM00001; EGF_like; 1.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 5.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 8.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 704 AA; 78137 MW; D47D5A30D5E42932 CRC64;

Query Match 27.3%; Score 686.5; DB 13; Length 704;
Best Local Similarity 30.0%; Pred. No. 2.5e-61;
Matches 164; Conservative 66; Mismatches 164; Indels 153; Gaps 22;

QY 23 GSASPDSEPDST-ECTDGYEWDPSQCHRDVNECLTIPACKGEMKCINHYGYLCL 81
Db 185 GPCSQQCRDTGSGYSCFVGQLOQDPGVNCDINECITGTHSCGIGQTCVNTLGSFRC- 243
QY 82 PRSAAVINDLHGGPPPPVPPAQHPNCPGPEPDDQDSCVDVDECAQALHDCRPSQDCH 141
Db 244 -----QDRTSCGTGYELTDDSRCKDIDECETGTHNCPDFICQ 281
QY 142 NLPGSYQC-----T 150
Db 282 NTPGSEFCRPLQCMNGFIQDALGNICIDINECLSTNMPGAGQICINTDGSYTCRISPS 341
QY 151 CPDGY--RKIGPECVDIDEC----- 168
Db 342 CGRGYHLNEDGTCDVDDECSSDQPCGEGHVCINGPNYRCEKSGYSFDVLSRTCIDI 401
QY 169 ----RY--RYCOHRCVNLPGSFRCPQEPQLGNPNRNSVDVNECDMGAPCEQRCFNSY 222
Db 402 NECRRYPGRGLCAKCENTPGSYCTCTMGFKLSSDGRSCEDLNECE--SSPCSECANVYG 460
QY 223 TFLCRCHQGYELHR-DGFCSDIDECY--SSYLQYRCVNEPGRFSCHPQ-GYQLLA- 277
Db 461 SYQCYCRRGFQSLDIDGISCEDIDECALPTGGHTICFRCINIPGSGFQCTCPSTGYRLAPN 520
```

```

QY 278 TRLCQDIDECESGAHOCSEATQCVNTHGGYRCVDTNRCVEPIQVSEN-----RCL 328
Db 521 ARNCQDIDECVAETHCSFNCFNIQGGFRCLSL-ECPENYRKSGDVTBRLKTDTRI 579
QY 329 --CPASNPCLREOP-SSIVHYWTITSERSV--PADVFQIQ-TSVVPG--AYNAFOIRA 380
Db 580 KSCRPNVNCVLPDVHTISHTVISLFTFRETFPEEILFLRAITPYPANQADLIIDITE 639
QY 381 GNSQGFYI--ROINNVMFAMLV-LARPVTGPREYVLDLEM-VTMNSLMSYRASSVRLTV 436
Db 640 GNLRESFDIILKRYMDGTGVQVRPIVGPFFHAILKLENNYVMGGVVSHR--NIVNVHI 697
QY 437 FVGAYTF 443
Db 698 FVSEYWF 704

RESULT 8
O42182 PRELIMINARY; PRT; 681 AA.
AC O42182;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FIBULIN-1 D.
GN FBLN1.
OS Brachydanio rerio (Zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
SEQUENCE FROM N.A.
RA Zhang H.-Y., Lardelli M., Ekblom P.;
RT "Sequence of zebrafish fibulin-1 and its expression in developing
RT heart and other embryonic organs.";
RL Dev. Genes Evol. 0:0-0(1997);
DR EMBL; AF013751; AB80944.1; -.
DR HSSP; P35555; IEMN.
DR ZFIN; ZDB-GENE-990415-73; fbln1.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF00008; EGF; 4.
DR SMART; SM00104; ANATO; 2.
DR SMART; SM00179; EGF_CA; 5.
DR SMART; SM00001; EGF_like; 4.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 6.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 681 AA; 74459 MW; 175C966305A46699 CRC64;

Query Match 26.0%; Score 653.5; DB 13; Length 681;
Best Local Similarity 28.7%; Pred. No. 5 6e-58;
Matches 155; Conservative 67; Mismatches 146; Indels 173; Gaps 20;

QY 39 CTGDEYWDPSQHRDNECLTIPEACKGEMKCNHYGGLCLPRSAVINDLHGEGPPP 98
Db 178 CLDGFKLTKDGKHCEDINECLLPHHCVTGERCINTLGSYRC----- 219
QY 99 PVPPAQHPNCPGYPDDQDSVDVDECAQALHDCRPSQDCHNLPGSYQC----- 149
Db 220 -----QREISCGTGYELTDNNKCKDIECDLGLTHNCAAMEQNTAGSFCRPRMCOAG 274
QY 150 -----TCPDGY----- 155
Db 275 FIQDALGSCIDINECVSVTALSRGQMCFTVGSFICQRHSVTGCRGYHLNAEGTRCVDID 334

```

```

QY 156 -----RKIGPECVDIDECRY-----RYCOHRCVNLP 181
Db 335 ECAGPNSCDHGHCINLVSYRCCERTGTFNFSITSRSCEIDECRNYPGRCAHKCNIL 394
QY 182 GSFRCECEPGFQGNRSCVDYNECDMGAPCEORCFNSYGTFLCRCHOGYEL-HRDGFS 240
Db 395 GSYKSCCTAGFKLADGRNCDNDNECE-SSPCSQGCANVYGSYQSRRYQLSDADGIT 453
QY 241 CSIDIDECSY--SSYLQYRCVNEPGRFSCHCP-QGYOLLAT-RLCQDIDECESGAHOCSE 296
Db 454 CEDIDECALPTGGHICSYRCHNTPGSEFHCCTPASGYTLAANGRSCQDIDELCTGTHSCSE 513
QY 297 AQTGVNPHGGYRCV-----DTNRCVEPIQVSENRCCLPASNPL 335
Db 514 SESCFTIQGGFRCLSFDCPANYRRSGDTRPRVDRADIIRC VKS-CQHNDISCVL---NPI 569
QY 336 CREQPSIVHYWTIT-----SERSVPAD-VFQIQANSVYPGAYNA-----FOIRAGNSQGD 386
Db 570 -----LSHSTAIISLPTTFREFNKPEEIVFLRSPTTHLPHMDSPEIVYDILEGNIONS 622
QY 387 F-YIROINN--VFAMLVLARPVTPGPREYVLDLEM-VTMNSLMSYRASSVRLTVFVGAYT 442
Db 623 FDIIRLDHGMIVGVVQVPRPLVGPVTVLKLAMNYVTNGVVSHR--NIINRVIYSEFW 680
QY 443 F 443
Db 681 F 681

RESULT 9
Q9HBQ5 PRELIMINARY; PRT; 495 AA.
AC Q9HBQ5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 54.3 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217999; AA017241.1; -.
DR HSSP; P35555; IEMN.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00181; EGF; 9.
DR SMART; SM00179; EGF_CA; 9.
DR SMART; SM00001; EGF_like; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
KW EGF-like domain; Glycoprotein; Hydroxylation; Hypothetical protein.
SQ SEQUENCE 495 AA; 54340 MW; C40434B6C82E3D70 CRC64;

Query Match 25.8%; Score 649; DB 4; Length 495;
Best Local Similarity 39.5%; Pred. No. 1.1e-57;
Matches 137; Conservative 44; Mismatches 128; Indels 38; Gaps 14;

QY 38 ECTDGYEWDPSQHRDNECLTIPEACKGEMKCNHYGGLCLPRSAVINDLHGEGPP 97
Db 163 QCKSGFTQDA-LGNCIDINECLSIAPCPICHTCINTEGSYTC----- 204
QY 98 PPVPPAQHPNCPGYPDDQDS-CVDVDECAQALHDCRPSQDCHNLPGSYQCTCPDGY- 155

```

```
Db 205 ---QKNVNP-CGRGYHLNEEGTRCVDVDECAAPAEPCGGRHRCVNSPGSFRCECKTGY 259
QY 156 -RKIGPECVIDEC-RY--RYCQHRVNLPGSFRCEQEPFOLGPNRRSCVDVNECDMGA 211
Db 260 FDGISRMCVGNQCORYPGRLCGHKCBNTLGSYLCSGVSFRLSVDRSCEDINECS-SS 318
QY 212 PCEQRCFNSYGTFLCRCHQGYELHR-DGFSCSDIDECY--SSYLCOYRCVNEPGRFSCH 268
Db 319 PCSQECANVYGSYQCYRRGYQLSDVGVTCEDIDECALPTGGHICSYRCINIPGSFQCS 378
QY 269 CP-QGYQLLAT-RLCQDIDECESGAHOCSEAQTCVNFHGGYRCVDTNRCVPEYIQVSEN 326
Db 379 CPSSGYRLAPNGRCQDIDECVTGIHNCINETCFNIQGGFRCL-APECPENYRRSAATR 437
QY 327 C--LCPASNPLCRPSSIVHRYWTITRSRVPADVFOICQATVYPG 371
Db 438 CERLPCHENRECSKPLRIYYHLSPFTNIQAPAVFRMGPSNAVPG 484

RESULT 10
Q9UH16 PRELIMINARY; PRT; 554 AA.
AC Q9UH16
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE DJ162H14.1 (FIBULIN 1) (FRAGMENT).
GN FBLN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98047; CAB62995.1; -.
DR HSSP; P35555; IEMN.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 3.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00104; ANATO; 3.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_like; 5.
DR PROSITE; PS01177; ANAPHYLATOXIN.1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN.2; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF.2; 3.
KW EGF-like domain; Glycoprotein; Hydroxylation.
FT NON_TER
SQ SEQUENCE 554 AA; 59767 MW; FE285184599A2982 CRC64;
```

```
Query Match 22.3%; Score 560; DB 4; Length 554;
Best Local Similarity 39.9%; Pred. No. 1.5e-48;
Matches 117; Conservative 33; Mismatches 107; Indels 36; Gaps 13;

QY 38 ECTDGYEWDPSQHCVDVNECLTPEACKGEMKINHYGYLCLPRSAAVINDLHGEGPP 97
Db 266 QCKSGFTQDA-LGNCIGINECLSLSAFCPTGHTCINTEGTYC----- 307
QY 98 PVPPAHPNCPGPEYFPDDDS--CYDVEDCAQALHDCRPSQDCHNLPGSYQCTCPDGY- 155
Db 308 ---QKNVNP-CGRGYHLNEEGTRCVGVDECAAPAEPCGGRHRCVNSPGSFRCECKTGY 362
QY 156 -RKIGPECVIDEC-RY--RYCQHRVNLPGSFRCEQEPFOLGPNRRSCVDVNECDMGA 211
Db 363 FDGISRMCVGNQCORYPGRLCGHKCBNTLGSYLCSGVSFRLSVDRSCEDINECS-SS 421
```

```
QY 212 PCEQRCFNSYGTFLCRCHQGYELHR-DGFSCSDIDECY--SSYLCOYRCVNEPGRFSCH 268
Db 422 PCSQECANVYGSYQCYRRGYQLSDVGVTCEDIDECALPTGGHICSYRCINIPGSFQCS 481
QY 269 CP-QGYQLLAT-RLCQDIDECESGAHOCSEAQTCVNFHGGYRCVDTNRCVPEY 319
Db 482 CPSSGYRLAPNGRCQDIDECVTGIHNCINETCFNIQGGFRCL-APECPENY 533

RESULT 11
Q95N23 PRELIMINARY; PRT; 689 AA.
AC Q95N23;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE F56H11.1B PROTEIN.
GN F56H11.1B.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68219; CAC35827.1; -.
DR EMBL; Z68749; CAC35818.1; -.
DR EMBL; Z68749; CAC35818.1; -.
DR EMBL; Z68219; CAC35818.1; JOINED.
SQ SEQUENCE 689 AA; 74625 MW; OFED2023E11D6AFB CRC64;

Query Match 22.1%; Score 555.5; DB 5; Length 689;
Best Local Similarity 26.4%; Pred. No. 5.7e-48;
Matches 144; Conservative 67; Mismatches 158; Indels 177; Gaps 22;

QY 39 CTDGYEWDPSQHCVDVNECLTPEACKGEMKINHYGYLCLPRSAAVINDLHGEGPP 98
Db 180 CRSGFDLAPDGMACVDIDECATLMDCLDSQRCINTPFSFKCI-RTLS----- 226
QY 99 PVPPAHPNCPGPEYFPDDQ--DSCVDVEDCAQALHDCRPSQDCHNLPGSYQCT----- 149
Db 227 -----CGTGYAMDSETERCDVDNGLSHDCCGLYQCRNTQGSYRCDAKCKGCG 276

QY 150 -----TCPDGYRKIGPECVIDEC-----RYRYC 173
Db 277 ELONPMTGCTSTICPNYYTPKNGMCNDIDECVTGHNCGAGECVNTPGSRFCQQKGNLC 336
QY 174 QH-----RCVNLP----- 181
Db 337 AHGEVNGATGFCEDVNECQGVGSGMECLNPGTYKCGPGYERNDAKKRCEDVDECI 396
QY 182 -----GSFRCEQEPFOLGPNRRSCVDVNECDMG--APCEQRCFNSYGT 223
Db 397 KFAGHVCDLSAECINTIGSFCECKKPGFQLADGRRCEDVNECTGTGAACEQCVNIPGS 456
QY 224 FLCRCHQGYELHRDGFSCSDIDECY----SSYLCOYRCVNEPGRFSCHQCYQLLAT- 278
Db 457 YQICDRGFALGPDGTGKCEDIDECSTWAGSNDLCMGCCINTKGSYLCCQCPGKTKPDG 516
QY 279 RLCQDIDECESGAHOCSEA-QTCVNFHGGYRCVDTNRCVPEYIQVSEN-----CL- 328
```

Db 517 RTCDVDECAMG--ECAGSDKVCNTLGSFKCHSID-CPTNYIHDSLNNQIADGVSCIK 573  
QY 329 -CPASNPLC-REQPSSIVHYRMTITSERSV--PADVFOIQATSVYPGA--YNAFQIRAGN 382  
Db 574 VCSTEDTEGLNHTREVLYQFRAVPSLKIISPIEVSRIVTHMGVPFSDVNDLYV--- 629  
QY 383 SQGDFYIRQINNVFAMVLARPVTGPREVLDLEMTMNSLSYRASSVLR-----LTVF 437  
Db 630 GORHFRIVOERNI-GIVQLVKPISGP-----TVETIKVNIHTKSRGTGVLAFNEALIEIS 683  
QY 438 VGAYTF 443  
Db 684 VSKYPF 689  
RESULT 12  
Q9T2S1 PRELIMINARY; PRT; 589 AA.  
AC Q9T2S1;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE FIBULIN-1D (FRAGMENT).  
GN FBLN1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoida;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CB1489 HIM-8 (E1489);  
RX MEDLINE=99120531; PubMed=9923656;  
RA Barth J.L., Argraves K.M., Roark E.F., Little C.D., Argraves W.S.;  
RT "Identification of chicken and C. elegans fibulin-1 homologs and  
characterization of the C. elegans fibulin-1 gene";  
RL Matrix Biol. 17:635-646(1998).  
DR EMBL; AF070477; AAC24035.1; -;  
DR HSSP; Pf16109; IFSB.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR Pfam; PF00008; EGF; 6.  
DR SMART; SM00179; EGF\_Ca; 5.  
DR SMART; SM00001; EGF\_like; 4.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS01186; EGF\_2; 5.  
DR PROSITE; PS01187; EGF\_Ca; 7.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
FT NON\_TER 1  
SQ SEQUENCE 589 AA; 63984 MW; 8EA3E8FCE0B97BE6 CRC64;  
Query Match 22.1%; Score 555; DB 5; Length 589;  
Best Local Similarity 29.1%; Pred. No. 5,3e-48;  
Matches 144; Conservative 68; Mismatches 162; Indels 120; Gaps 25;  
QY 39 CTDGYEWDPSQHRDVRNECLTIPEACKGEMKCNHYGYLCLPRSAVINDLHGEGPPP 98  
Db 127 CGTGYAMDSETERCDVRDECLGSHDGCPLPYQCRNTQGSYRCDAKCK-----GDG--- 176  
QY 99 PVPPAQHPN-----CPPGYEPDQSDVDVDECAQALHDCRPSQDCHNLPGSYQC 149  
Db 177 -----ELQNPMTGECTSITCPNGYKPN-GMCDIDECVTG-HNCGAGECVNTPGSFRC 229  
QY 150 -----TCDGYR-----KIGP-----EC 162  
Db 230 QQKGNLCAGYEVNATGTGCDVNECQQVCGSGMECINLPGTYYKCKGPGYFNDAKKRC 289  
QY 163 VDIDECRYCQH-----RCVNLPGSFRQCQEPGFGQLPNRRSCVDVNECDMG-APCEQ 215  
Db 290 EDVDEC-IFAGHVCDLSACINTIGSFCKCKPGFQLASDGRCEDVNECTTGIAACEQ 348

QY 216 RCENSYGTFLCRCHQGYELHRDGFSCSDIDECY-----SSYLCOYRCVNEPGRFSCHCPQ 271  
Db 349 KCVNIYSGQCICDRGFALGPDGTCKEDIDECISWAGSNDLCMGCCINTKGYLQCCPP 408  
QY 272 GYOLLAT-RLCODIDECSEGAHOCSEA-QTCVNFHGYRCVDNRCVPEVQVSENK--- 326  
Db 409 GYKIQPDGRTCDVDDECAMG--ECAGSDKVCNTLGSFKCHSID-CPTNYIHDSLNNQI 465  
QY 327 -----CL--CPASNPLC-REQPSSIVHYRMTITSERSV--PADVFOIQATSVYPGA--YN 374  
Db 466 ADGYSCIKVCSTEDTEGLNHTREVLYQFRAVPSLKIISPIEVSRIVTHMGVPFSDVYN 525  
QY 375 AFQIRAGNSQGFYIRQINNVFAMVLARPVTGPREVLDLEMTMNSLSYRASSVLR- 433  
Db 526 LDYV-----GORHFRIVOERNI-GIVQLVKPISGP-----TVETIKVNIHTKSRGTGVLAF 575  
QY 434 ----LTVFVGAYTF 443  
Db 576 NEALIEISVSKYPF 589  
RESULT 13  
O18026 PRELIMINARY; PRT; 798 AA.  
AC O18026; Q20903;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE F56H11.1 PROTEIN.  
GN F56H11.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoida;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wilkinson J.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,  
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,  
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans.";  
RL Nature 368:32-38(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Lloyd C.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z68749; CAA92962.1; -;  
DR EMBL; Z68219; CAA92962.1; JOINED.  
DR EMBL; Z68219; CAA92483.1; -;  
DR EMBL; Z68749; CAA92483.1; JOINED.  
DR HSSP; Pf16109; IFSB.  
DR InterPro; IPR000020; Anaphylatoxin.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR Pfam; PF01821; ANATO; 2.  
DR Pfam; PF00008; EGF; 5.  
DR SMART; SM00104; ANATO; 2.  
DR SMART; SM00179; EGF\_Ca; 4.  
DR SMART; SM00001; EGF\_like; 6.  
DR PROSITE; PS01177; ANAPHYLATOXIN\_1; UNKNOWN\_1.

```
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 8.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 798 AA; 87205 MW; 3BFEE9ED54D8BF9 CRC64;

Query Match      21.1%; Score 529.5; DB 5; Length 798;
Best Local Similarity 24.5%; Pred. No. 3e-45;
Matches 146; Conservative 61; Mismatches 154; Indels 235; Gaps 20;

QY 39 CTGDEWDPDSQHCRDVNECLTTPACKGEMKCIHYGYLCLPRSAAVINDLHGEGPPP 98
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 180 CRSGFDLADPGMACVDRNECLTRQSPCTQSEDVCNTIGGYICQRRLSLRVPHRHRANRIG 239
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 99 PVPPAQHPNP-----CPGPEPDDQSDVDDECAQALHDCRPSQDC 140
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 240 NAFRRMDDPYRAGEYREASQANTFEGCPMGW-LFQHGHCVDIDECALMDCLLESQRC 298
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 141 HNLPGSYQC---TCPDGYRK----- 157
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 299 LNTPGSEKCIRTLSCGTGYAMDSETERNNCFLLIINNTECNKYFFVEDVDECNLGSDDCG 358
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 158 -----IGPECV----- 163
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 359 PLYQCRNTQSGYRCDAKKCGDGLQNPMTGEYIDEQVTGHNCGAGECVNTPGSRFCQOK 418
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 164 -----DIDECRYRYC-QHRCVNLP----- 181
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 419 GNLCAGHYEVNGATGCEDEVNECQGVCGSMWCINLPGTYYKCKGPGYEENDAKKCEDV 478
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 182 -----GFRCECPGFGOLGPNRRNSCDVNECDMG-APCEORCFN 219
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 479 DECIFAGHVCDLSAECINTIGSECKKPGFQLASDGRRCEDVNECTTGIAACEQKCVN 538
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 220 SYGFLCRHQGYELHRDGFSCSDIDECVY-----SSVLCQYRCVNPGRFSCHCPQGYQL 275
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 539 IPGSIQCIDRGFALGPDGPKCEDIDECSTWAGSGNDLCMGGICNTKGSYLCOCPPGYKI 598
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 276 LAT-RLCQDIDECESGAHQSEA-QTCVNFHGGYRCVDTNRCVPEYIOVSENRLC---P 330
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 599 QPDGRTCDVDECCAMG--ECAGSDKVCVNTLGSFKCHSID-CPTNVIHDSLNKNRCNRQP 655
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 331 ASNPL---CREQPSIVHRYWTITSERSVP-----ADV---FQIQAT 366
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 656 SAGCLPECSKVPLFLYIQFISLA--RAVPISSHRPAITLFKVSAPNHADTEVNFELQK 713
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 367 SVYPGAYNA-----FOIRAGNSQGDYIRQINNPFAMLVLARPYTGPREYVLDL 415
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 714 TTIVGAPNVLPAIRANFLLOKGEKRN-----AVVTLRSLDGPQTVKLQL 759
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 14
O08999 PRELIMINARY; PRT; 1833 AA.
ID O08999
AC O08999;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE LATENT TGF-BETA BINDING PROTEIN-2.
GN LTBP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Fang J., Li X., Smiley E., Francke U., Mecham R.P., Bonadio J.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004874; AAB61611.1;
DR HSSP; P35555; 1EMN.
DR MGD; MGI:99502; Ltbp2.
DR InterPro; IPR002086; Aldehyde_dehydr.
```

```
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002212; TB.
DR Pfam; PF00008; EGF; 16.
DR Pfam; PF00683; TB; 4.
DR SMART; SM00179; EGF_CA; 16.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; 12.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 10.
DR PROSITE; PS01187; EGF_CA; 16.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 1833 AA; 197928 MW; 3FCD6BD31E9FC6E CRC64;

Query Match      20.8%; Score 523.5; DB 11; Length 1833;
Best Local Similarity 30.5%; Pred. No. 3.4e-44;
Matches 128; Conservative 44; Mismatches 139; Indels 109; Gaps 19;

QY 33 PDSY-TECTDGYEWDPDQSHCRDVNECLTTPACKGEMKCIHYGYLCLPRSAAVINDL 91
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 876 PNGRCVCSPGYQLHPSQDYCTDDNECMNP--CEGRGRCVNSVGSYSL----- 923
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 92 HGEPPPVPVPAQHNPCCPGY---EPDQDSCVDVDEC----- 127
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 924 -----CYPGYTLVLTGTQECQDIDECQEPGVCVSGGRCSNTEGSHCE 966
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 128 -----AQLHDCR-----PSQCHNLPGSYQC-TCPDGKRGKIGPECVDIDEC- 168
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 967 CDGTYIMVRKGCQDINECRHPTCTPDGRCVNSPGSYTCLACEGYVGSGSCVDVNECL 1026
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 169 RYRYCOH-RCVNLPGSFRQCCEPGFOLGPNRRNSCDVNECDMGAPCEQ-RCFNSYGTFLC 226
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1027 TPGICITGRCINMEGSRCSCEPGYEVTDPKKGCRDVDECASTRASCTGLCLNTEGFTC 1086
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 227 -RCHQGYELHRDGFSCSDIDECVSYSLCQYRCVNPGRFSC-HCPQGYQ--LLATRLCQ 282
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1087 SACQSYWYNEDGTACEDIDECAPGVCVCTGTNTVGSFCKDCQDGGYRPNPLGNR-CE 1145
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 283 DIDECSGAHQSEAQTCVNFHGGYRCV-----DTNRCVPEYIOVSENRLC 328
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1146 DVDECSGPOSSCRGGE-CKNTEGSIQCLCHQGFQLVNGTMCDVNECVGEEHCAHGECL 1204
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 329 CPASNPLCREQPSIVHRYWTITSERSVPA-DVFQIQATSVTPGAYNAFIRAGNSQGD 387
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1205 NSLGSFFCLCAPG-----FASAEGRTRCQDVDECAATDPCPGGH-----CVNTEGSF 1251
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 15
Q9WUH9 PRELIMINARY; PRT; 2906 AA.
ID Q9WUH9
AC Q9WUH9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FIBRILLIN-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99350231; PubMed=10419698;
RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashiwara N., Wallner E.,
RA Kanwar Y.S.;
RT "Cloning of rat fibrillin-2 cDNA and its role in branching
morphogenesis of embryonic lung."
RL Dev. Biol. 212:229-242(1999).
DR EMBL; AF135060; AAD34439.1;
DR HSSP; P35555; 1EMN.
DR InterPro; IPR002086; Aldehyde_dehydr.
```



```
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR002212; TB.
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR PRINTS: PR00010; EGF_BLOOD.
DR SMART: SM00179; EGF_CA; 42.
DR SMART: SM00001; EGF_like; 4.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
DR PROSITE: PS00010; ASX_HYDROXYL; 43.
DR PROSITE: PS00022; EGF_1; UNKNOWN_2.
DR PROSITE: PS01186; EGF_2; 36.
DR PROSITE: PS01187; EGF_CA; 43.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2906 AA; 313371 MW; 9EE64E727044EF58 CRC64;

Query Match      20.8%; Score 523; DB 11; Length 2906;
Best Local Similarity 33.4%; Pred. No. 6.8e-44;
Matches 113; Conservative 38; Mismatches 109; Indels 78; Gaps 15;

QY 38 ECTDGYEWDPSOCHRDVNECLTPEACKGEMKCNHYGYLCIPRSAVINDLHGEGPP 97
Db 1176 DCPIGHLSPSREDCIDNECSLSDNLCRNG-KCVNMIGTYQC----- 1217

QY 98 PVVPPAQHPNPPGGE-PDQDQSDVDVDECAQALHDCRPQSDCHNLPGSYQCTCPDGYR 156
Db 1218 -----SCNPGYQATPDROGCSIDIDECIMNGGC--DTQCTNSEGSEYECSCSEGYA 1265

QY 157 KI--GPECVDIDECYR--YQ--HRCVNLPGSPROCEPGFQLGPNRSCVDVNECD--- 208
Db 1266 LMPDGRSCADIDECENNPIDICDGGQCTNIPGYRCLCYDGFMA5MDMKTCIDVNECDLNP 1325

QY 209 ---MGAPCEQRCFNSYGTFLCRCHQGYELHRDGFSCSDIDECSSYLYQYR--CVNEPG 263
Db 1326 NICMFECE---NTKGSFICHQLGY5VKKGATGCTDVECEIGAHCNDMHASCLNVP 1381

QY 264 RFSCHPQGYQLLATRLCQDIDECESGAHQCSAQTCVNFPHGGYRCV----- 310
Db 1382 SFKCSREGWVGNGIK-CIDLDECANGTHQCSINAQCNTPGSYRCACSGEGTGDGFTCS 1440

QY 311 DTRNCEVEPYIQVSEN-----RCLC-----PASN 333
Db 1441 DVDECAE-NINLCENGQCLNVPGAYRCECEMGFTPASD 1477
```

Search completed: September 25, 2002, 09:46:31  
Job time: 256 sec

